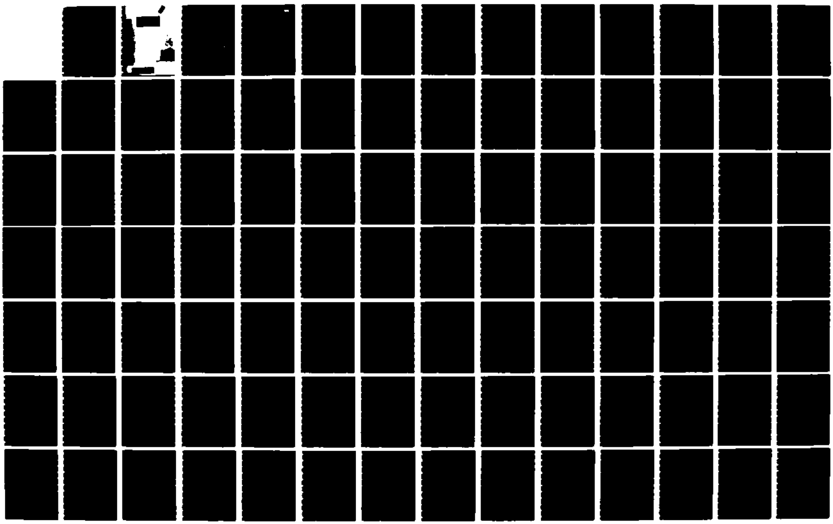
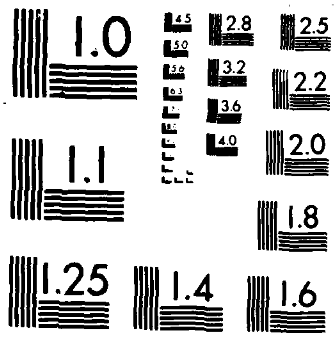
AD-A174 522 FUNCTIONAL DESCRIPTION FOR A REPLACEMENT INPUT SYSTEM 1/2
(U) LOGISTICS MANAGEMENT INST BETHESDA MD
D F EGAN ET AL JUL 86 LMI-DL504 MDA903-85-C-0139
UNCLASSIFIED F/G 5/2 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(1)

FUNCTIONAL DESCRIPTION FOR A REPLACEMENT INPUT SYSTEM

July 1986

Donald F. Egan
Richard W. Hartt
Philip J. Tarnoff

Prepared pursuant to Department of Defense Contract No. MDA903-85-C-0139 (Task Order DL504). The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official Department of Defense position, policy, or decision, unless designated by other official documentation.

LOGISTICS MANAGEMENT INSTITUTE
6400 Goldsboro Road
Bethesda, Maryland 20817-5886

Executive Summary

FUNCTIONAL DESCRIPTION FOR A REPLACEMENT INPUT SYSTEM

The Defense Technical Information Center (DTIC) processes a large and growing volume of documents. It adds more than 48,000 new documents a year to its data bases and processes almost as many change transactions. In the past 5 years DTIC has experienced a 10 percent increase in input workload with a decrease in staff. The current input software – primarily batch oriented and difficult to maintain – precludes significant productivity improvement. DTIC needs a new way to process its input.

We recommend a replacement input system detailed in this functional description. It features:

- Reduced document processing cost
- Reduced time for document microfilming and entry into the data bases
- Improved accuracy of data
- Reduced effort to maintain or revise the system.

The functional description begins with the current methods and procedures and from these develops the requirements for a replacement system. It lays out the recommended system and presents configurations for meeting those requirements which will operate with other DTIC hardware and software.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

SECTION 1. GENERAL

1.1 Purpose.

This functional description (FD) for the Defense Technical Information Center (DTIC) Input System Replacement project provides:

- Descriptions of the organizational functions and processes to be supported by the proposed system
- The system requirements to be satisfied, which will serve as a basis for a common understanding between the user and the developer
- Descriptions of three alternative approaches to meeting those requirements.

1.2 Project References.

The development of this document is authorized by Logistics Management Institute (LMI) Task Order No. DL504 (MDA0166-165), "DTIC Input System Requirements," undertaken by LMI at the request of the Defense Logistics Agency (DLA). The project is sponsored by the Office of Information Systems and Technology of DTIC.

The information in this FD was derived primarily from interviews with DTIC staff members (see Table 1-1) between 1 May and 3 July 1985, and the descriptions apply to the methods and procedures then in effect. The interviews were conducted jointly by members of the Information Systems Division and LMI. Interviews concentrated on the three directorates performing functions related to document input or operating the existing input system: Data Base Services (DTIC-H), Document Services (DTIC-F), and Telecommunications and Automated Data Processing (ADP) Systems (DTIC-Z).

In addition to operating an in-house facility for document input, DTIC sponsors and supports the Shared Bibliographic Input Network (SBIN) for direct entry of

TABLE 1-1. DTIC INTERVIEWS

ORGANIZATION	NUMBER OF PERSONS INTERVIEWED
Data Base Services Directorate (H)	3
Descriptive Cataloging	7
Information Analysis	3
Technical Reports	6
Management Information	6
Data Base Output	2
Document Services Directorate (F)	3
Acquisition	5
Selection	4
Reference Services	1
Telecommunications and Automated Data Processing (ADP) Systems Directorate (Z)	2
Systems Design	3 plus 1 contractor
Technical Support	4
Office of Planning and Management	2
Office of Information Systems and Technology	3
Total	54 + 1

document citations by remote technical libraries and information analysis centers (IACs) of the Department of Defense (DoD). A survey questionnaire was mailed to 54 SBIN members to help determine SBIN requirements for the input system replacement. The study team also visited nine SBIN sites and six IACs to gather additional information about requirements. The SBIN sites and IACs visited are listed in Appendix B.

The following references were reviewed as part of the requirements determination and in the development of this report:

- LMI Task Order No. DL504, "DTIC Input System Requirements," Logistics Management Institute, 15 February 1985.
- "Local Automation Model: Functional Description," Logistics Management Institute, September 1983.
- "Department of Defense Automated Data Systems Documentation Standards," DoD Standard 7935.1-S, 13 September 1977.
- "DTIC Long Range ADP and Telecommunications Plan," Applied Management Systems, Inc., July 1984.
- "Technical Reports Input Systems Analysis Program Specification Working Draft," Advanced Technology Company, July 1985.
- "Defense Technical Information Center Cataloging Guidelines," DTIC TR84/1, January 1984.
- "Research and Technology Work Unit Information System Manual," DoD Standard 3200.12-M-1, August 1984.
- "Independent Research and Development Data Bank Input Manual," *DLAM 4185.9, January 1981.*
- "Input Into the Defense Technical Information Center Technical Data Base," Office of the Inspector General Report No. 84-007, 8 November 1983.
- "DTIC 2000 – A Corporate Plan for the Future," DTIC/TR-84/3, July 1984.

1.3 Terms and Abbreviations.

The following terms and abbreviations are used in this document:

- AD Number A seven-character alphanumeric identifier preceded by "AD" that uniquely identifies every document in DTIC.
- AQ The DTIC Acquisitions data base.
- COSATI Committee on Scientific and Technical Information.
- DLA Defense Logistics Agency, DTIC's parent organization.
- DRIT DTIC Retrieval and Indexing Terminology, a controlled-subject vocabulary used for TR, WUIS, and IR&D documents.
- DROLS Defense Research, Development, Text and Evaluation (RDT&E) On-line System, a computerized document input and retrieval system used for TR, WUIS, and IR&D data bases.

- <u>DTIC</u> Defense Technical Information Center.
- <u>IAC</u> Information Analysis Center.
- <u>IR&D</u> Independent Research and Development Data Base, a data base of independent research being conducted by DoD-sponsored contractors.
- <u>LAM</u> Local Automation Model, a DTIC project to combine SBIN input and local library services for remote libraries.
- <u>MAD</u> Master AD, a DTIC computer file of more than 1.2 million complete references to the DTIC technical reports collection; more commonly referred to as the TR data base.
- <u>MAI</u> Machine Aided Indexing, a series of computer programs that determine applicable DRIT subject terms automatically, the basis of information in document abstracts.
- <u>MINI-MAD</u> Miniature MAD, DTIC computer files (there may be several MINI-MADs at any one time) consisting of current transactions to be merged into the MAD.
- <u>NSA</u> National Security Agency.
- <u>OCR</u> Optical character recognition.
- <u>Pipeline</u> The flow of TRs through various DTIC organizations, which process them for input into DROLS, TAB, and other announcement products.
- <u>SBIN</u> Shared Bibliographic Input Network, a process by which DoD technical libraries and IACs input TR citations directly to the TR data base.
- <u>RTIS</u> Remote Terminal Input Subsystem, a subsystem of DROLS used for TR, WUIS, and IR&D data base inputs.
- <u>TAB</u> Technical Abstract Bulletin, a DTIC product that announces recently received TRs at approximately two-week intervals (TAB cycles).
- <u>TR</u> Technical Report.
- <u>WUIS</u> Work Unit Information System, a DTIC data base for maintaining status of DoD research projects.

SECTION 2. SYSTEM SUMMARY

2.1 Background.

DTIC serves as the clearinghouse for technical and scientific information within DoD. DTIC operates the Defense Research, Development, Testing, and Evaluation (RDT&E) On-Line System (DROLS), which provides access to the Technical Reports (TR) data base of information on more than 1.5 million TRs prepared or sponsored by DoD organizations.

Similarly, DROLS offers access to the Work Unit Information System (WUIS) and the Independent Research and Development (IR&D) data base, which include information about DoD-sponsored and contractor-sponsored research projects. In addition to these three data bases, DTIC is also developing a news data base of DoD Program Element Descriptions (PEDs) to provide information about RDT&E funding requests throughout DoD. In addition to these user data bases, DTIC operates support data bases such as the Acquisitions (AQ) data base used to monitor requests for submission of TRs to DTIC.

DTIC also supports the development and operation of programs and automated systems designed expressly to enhance information-sharing throughout the DoD scientific and technical community. For example, the SBIN supports direct cataloging of TRs by libraries participating in the program. This speeds the dissemination of technical and scientific information and broadens participation in the cataloging process to increase the resources available for information sharing. The Local Automation Model (LAM) project will make it easier to catalog directly into the TR data base in support of SBIN objectives.

The Remote Terminal Input Subsystem (RTIS) of DROLS now serves at DTIC to process the bulk of the inputs directed to the TR and IR&D data bases.

Approximately 48,000 new entries and 45,000 updates or corrections are submitted annually through the RTIS. Inputs to the WUIS data base (approximately 9,500 records a year) rely mainly on a separate input process; only a small portion is submitted through RTIS. RTIS supports preparation of data base inputs by DTIC staff members (catalogers, indexers, and data entry operators), as well as more than 50 SBIN-member libraries and Information Analysis Centers (IACs). Using a different data base methodology, the Acquisition system processes approximately 7,000 transactions a year.

Even with the volume of reports processed, DTIC does not catalog all DoD-supported reports. The DoD Inspector General, in an audit report published 8 November 1983 (Report No. 84-007), estimated that, over a 7-year period, DTIC had received only 55 percent of all scientific and technical documents derived from DoD in-house research.

DoD has taken steps to improve the flow of scientific and technical documents into DTIC for both cataloging and secondary distribution, and DTIC is increasing its acquisition efforts. To meet the expected increases in workload, DTIC requires an input system capable of supporting improved levels of staff productivity for cataloging and data entry. The system must also meet user expectations for ease of use and responsiveness so that shared cataloging remains a viable means of sharing the costs associated with timely dissemination of technical information.

In planning for replacement of the current input system, a full range of functions and processes associated with document inputs is considered. In terms of data bases, the definition of replacement system requirements covers inputs to the TR, WUIS, IR&D, and AQ data bases. Organizationally, the definition covers (within DTIC) the Data Base Input Services Division of the Data Base Services Directorate (DTIC-H) and the Acquisition and Selection Sections of the Document Services Directorate (DTIC-F). These organizations perform functions and processes ranging from document acquisition through providing a computer record ready to be processed by the various data base update and output programs. This functional description (FD) includes the manual procedures and the data processing system used to support their activities.

This FD also covers the data processing steps, computer operations, and auxiliary files associated with inputting documents and updating data bases. Data processing for document inputs and data base updates is the responsibility of the Telecommunications and ADP Systems Directorate (DTIC-Z). Computer operations include machine-aided indexing (MAI), edit-checks on all entered or revised citations, and file updating. Auxiliary files used for document input include the source header files, corporate author hierarchy and acronyms file, and the DTIC Retrieval and Indexing Terminology (DRIT) file.

Other continuing projects in DTIC may affect—or be affected by—replacement of the input system. The projects considered during requirements definition include: the software normalization project under way within the Telecommunications and ADP Systems Directorate; design and development of a production control system for the Document Services Directorate; application of optical character recognition (OCR) to cataloging, under review by the Office of Information Systems and Technology; assessment of data base requirements, under consideration by the Data Base Services Directorate; and the Data Base Services Directorate project to update the DRIT. The relationship among these projects and their influence on development of the replacement input system is described in this document.

2.2 Objectives.

The purpose of replacing the input system is to increase the flow of scientific and technical document citations and related information into DTIC data bases. This objective can be attained through a combination of improvements in automated system performance, reducing the time and effort required to enter accurate data into data bases and eliminating the barriers to shared cataloging. Major performance requirements for the replacement system include the following:

- User access
  - Simplify terminal operations and user access to the input system
  - Improve system response time
  - Improve access for SBIN and IAC users
- User productivity
  - Reduce or eliminate duplication of data entries
  - Eliminate key-punched inputs
  - Provide on-line data entry and real-time edits
  - Make auxiliary files available on-line
  - Improve MAI techniques
  - Perform real-time, on-line checking for duplicate citations
  - Prepare forms and letters automatically
- System operations
  - Improve hardware reliability
  - Improve software reliability and maintainability
  - Reduce manual intervention in the operation of the system
  - Simplify operation of the input system
  - Improve system documentation
- Management Support
  - Simplify report processing
  - Provide production statistics
- System configuration and growth
  - Accept a variety of input media (e.g., tape, diskette, OCR)
  - Accept batch transmissions of citations from remote sites
  - Reduce the hardware dependency of applications software

- Provide for growth of the system (number of users)
- Allow future modification or functional expansion of the system.

2.3 Existing Methods and Procedures.

2.3.1 Organizational Responsibilities.

DTIC is organized into three major directorates and several support offices (see Figure 2-1). This section describes the organizational activities associated with inputting documents and other RDT&E-related information into three DTIC data bases: TR, WUIS, and IR&D. Within DTIC, three directorates perform or support input functions and processes: the Document Services Directorate, the Data Base Services Directorate, and the Telecommunications and ADP Systems Directorate.

The Document Services Directorate (Figure 2-2) is responsible for handling the physical document collection, including maintenance of the inventory, microfilm processing, and order processing. With respect to document input, the Directorate performs the following functions: *mail receipt*, document acquisition, and document selection.

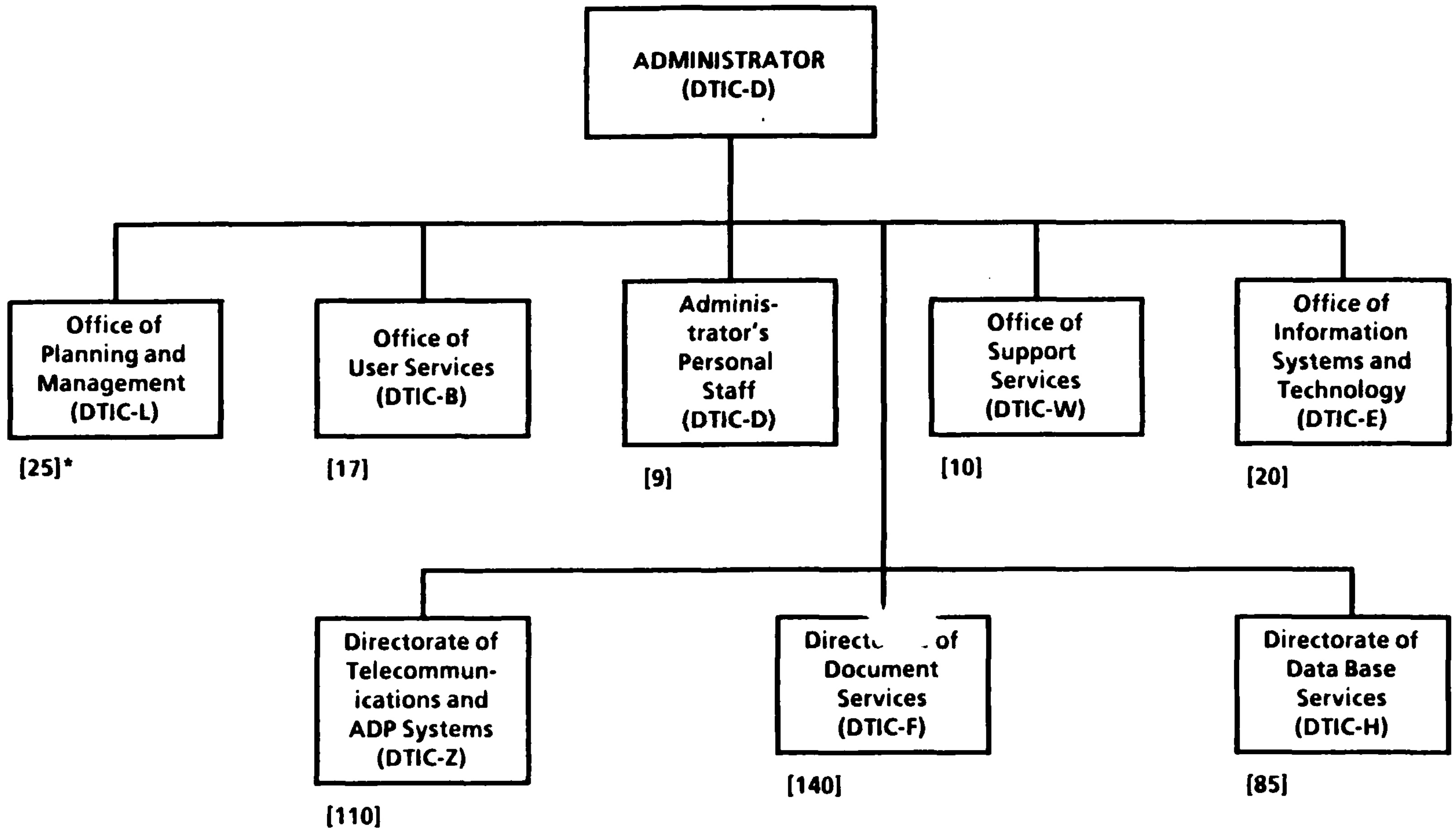
The Data Base Services Directorate (Figure 2-3) is responsible for maintaining the data in the data bases. It consists of the Data Base Management and the Analysis Divisions and involves four branches* that form the core input capability in DTIC:

- Bibliographic Data Base
- Subject Analysis – responsible for subject indexing
- Data Base Support – responsible for verification of TR document citations
- Research & Engineering (R&E) Programs Data Base – responsible for maintenance of the WUIS and IR&D data bases.

---

*The fifth branch: Retrieval Analysis Branch does not participate in input processes.

FIGURE 2-1. ORGANIZATION OF DTIC



*Indicates approximate number of employees or positions

FIGURE 2-2. DIRECTORATE OF DOCUMENT SERVICES (DTIC-F)

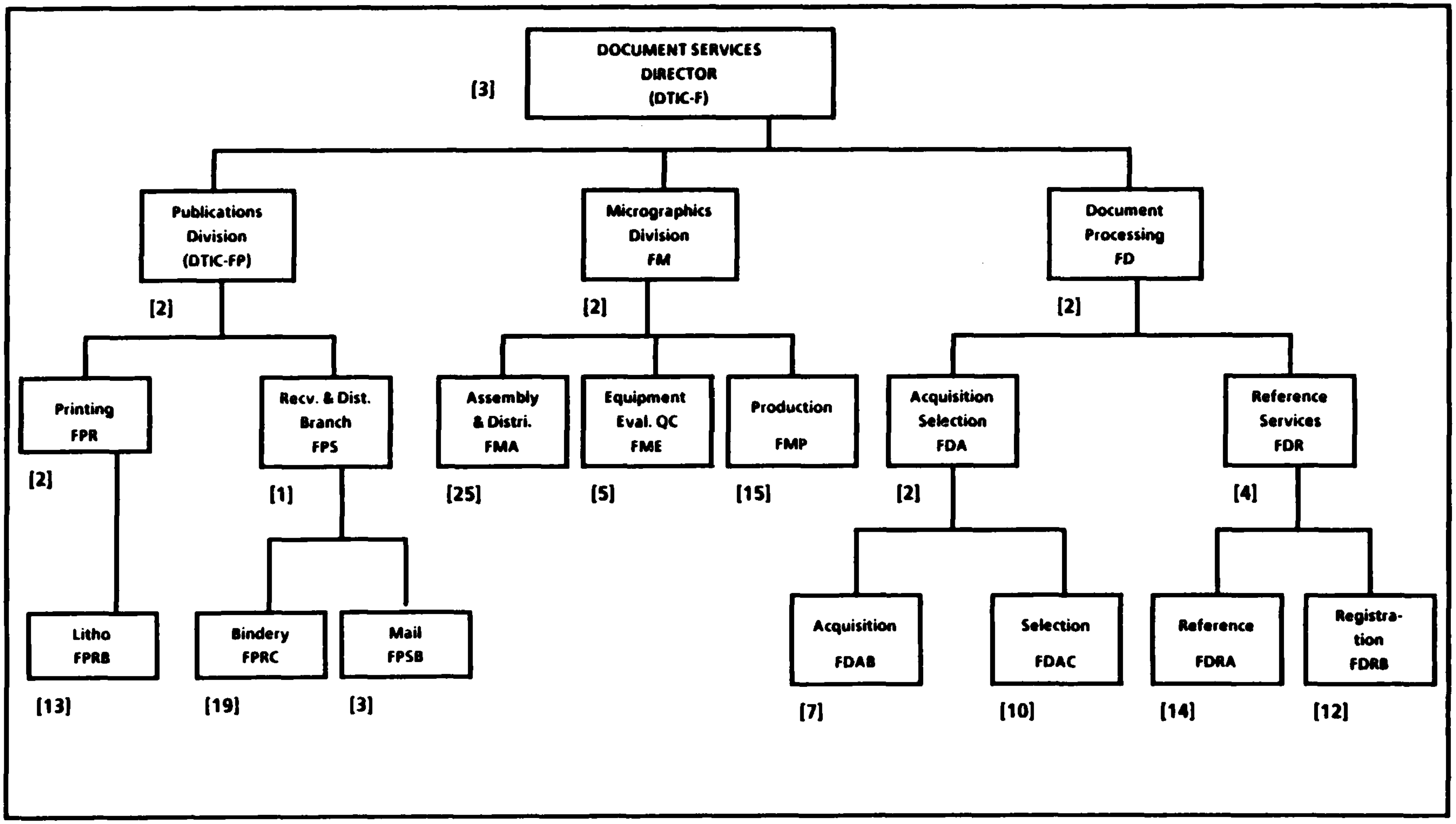
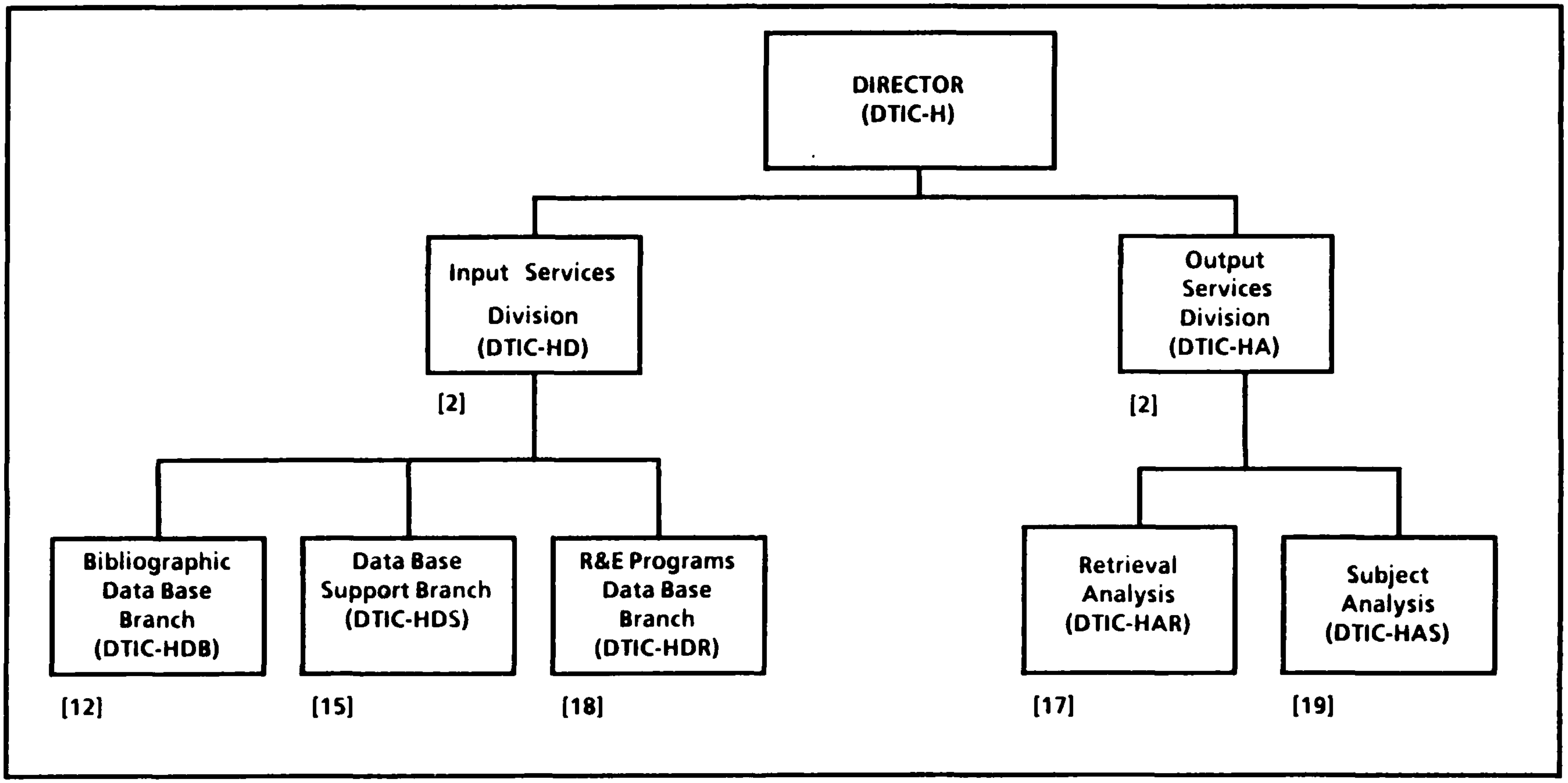


FIGURE 2-3. ORGANIZATION OF THE DIRECTORATE OF DATA BASE SERVICES (DTIC-H)



The process of passing TR input through these participating organizations is referred to as "the pipeline."

The Telecommunications and ADP Systems Directorate (Figure 2-4) is the data processing organization for DTIC, responsible for software development, maintenance of all production software, and computer operations.

2.3.2 Current Input Processing.

The discussion of present methods and procedures begins with the input of document citations into the TR data base. The description follows the sequence of events that are normally encountered when a new TR is added to the TR data base and traces a document as it moves through the organizations involved in document input. Figure 2-5 shows an overview of TR input processing by the DTIC staff. A more detailed flow chart of the processing steps involved is included in Appendix D.

The description of processing inputs for the TR data base concludes with a description of SBIN input processing in Section 2.3.9: Inputs to the WUIS and IR&D data bases are processed almost entirely by the R&E Program Data Base Branch (DTIC-HDR) (see Figure 2-3). Its operations are discussed in Section 2.3.10.

The processes described in the following pages reflect the organization as of 1 August 1985, when the analysis was completed. Since that time the Directorate of Data Base Services has undergone a major reorganization, which is only partially reflected in this document. The new organizational names and reporting relationships are reflected in the organization charts and in the section headings. The reorganization did not greatly affect the functional flow of the system, and, therefore, the text itself was not modified. Some of the changes are summarized in the following paragraph.

The Bibliographic Data Base Branch is now required to proof all of its documents, not just the classified documents. The Data Base Support group is now responsible for all TR subject and management data entry. This clerical support was

FIGURE 2-4. DIRECTORATE OF TELECOMMUNICATIONS AND ADP SERVICES (DTIC-Z)

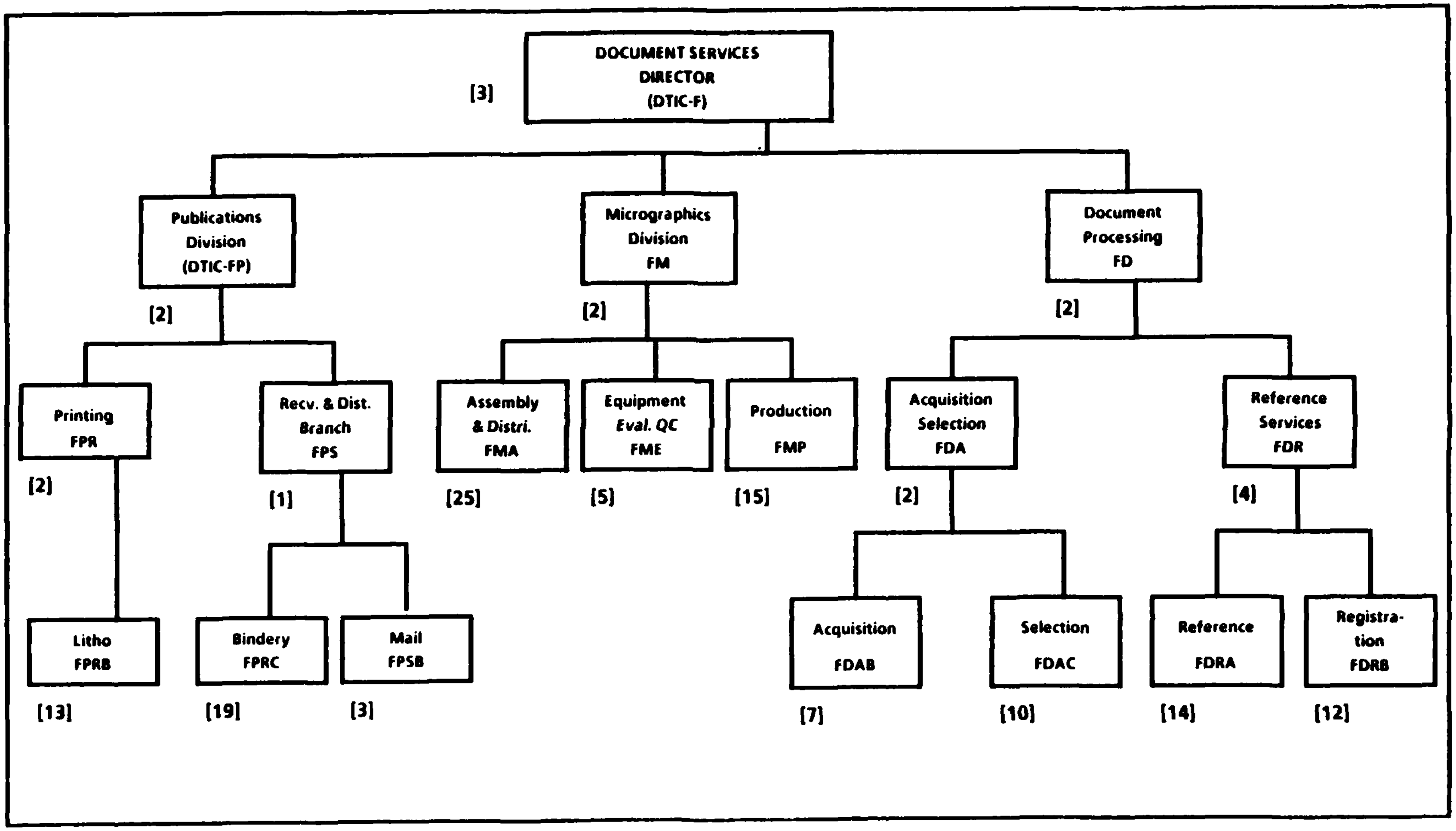


FIGURE 2-4. DIRECTORATE OF TELECOMMUNICATIONS AND ADP SERVICES (DTIC-Z)

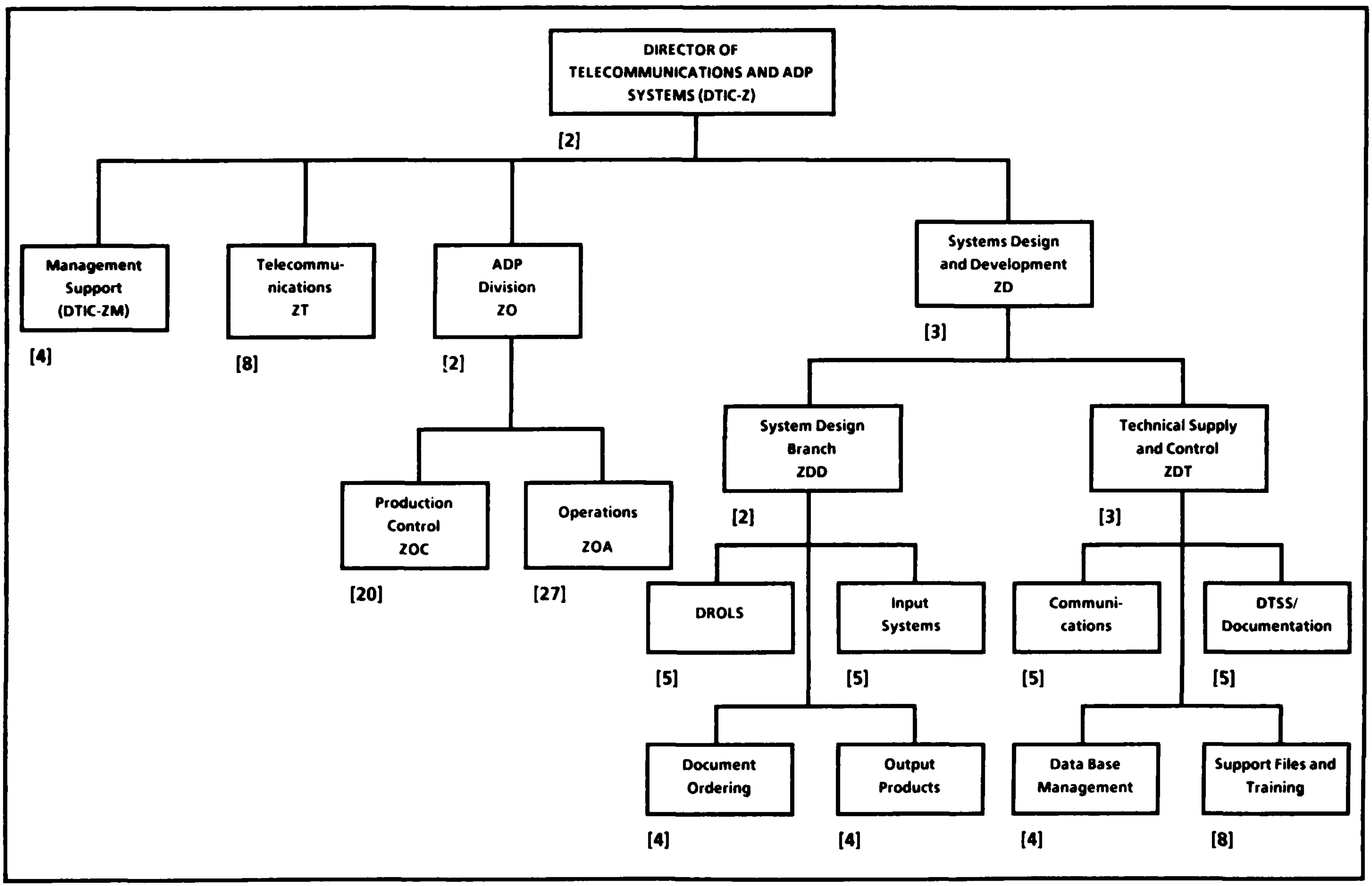
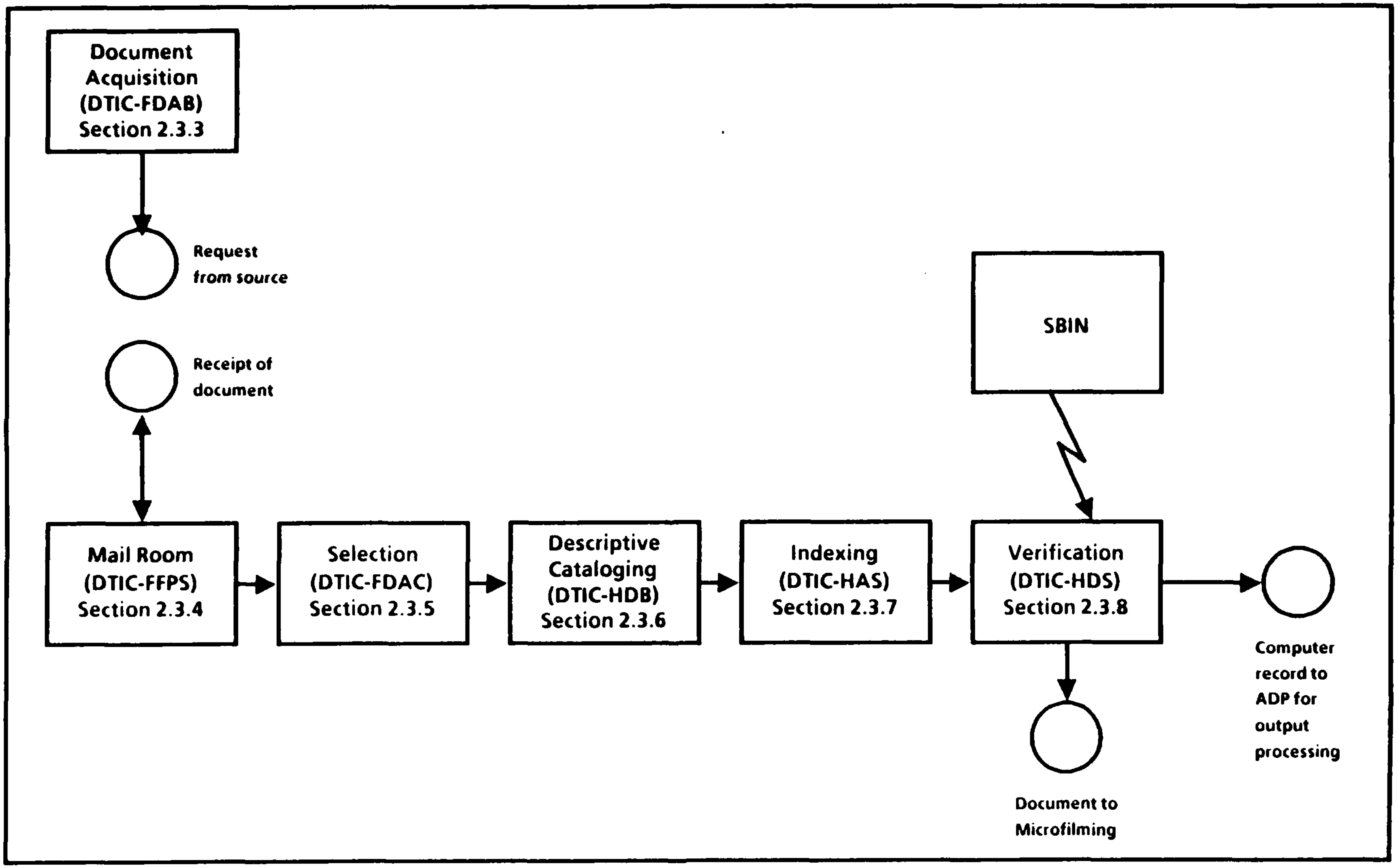


FIGURE 2-5. OVERVIEW OF TR INPUT PROCESSING DISPLAYING DTIC ORGANIZATIONS INVOLVED



transferred from the Subject Analysis and R&E Programs Data Base Branches. The Subject Analysis Branch has been transferred to another division, but its functions have not substantially changed.

2.3.3 Acquisition (DTIC-FDAB).

The Acquisition Section is responsible for obtaining new documents for the TR data base. Although information about potential acquisitions comes from a variety of sources, the three methods used most are reviews of in-house bibliographies of missing TR numbers, bibliographies from other source acquisitions, and requests from users received through the Reference Services Branch, and the "DD350" file. The DD350 file, a commercial data base containing information on DoD contracts, and the WUIS are compared with the TR for matching contract numbers. Whenever items are found in one of the data bases for which there is no corresponding item in the TR file, the source (i.e., organization identified on the data base record) is asked to get the document.

Once a decision is made to obtain a document, the Acquisition Section checks DROLS to find out whether the document is already in DTIC and checks the AQ data base to find out whether attempts have been made to get it. The AQ data base, which now comprises 12,000 records, is maintained on the Sperry 1100/61 computer, with the BASIS data base management system (DBMS). The AQ data base keeps basic bibliographic information about each document, such as author and title, and a history of efforts to acquire it, including the source of the document, the date requested, and the identity of the requester. If the document might have been added to the collection prior to 1982, a manual card file is also checked.

If the document is not found, a "Request for Scientific and Technical Report" form letter (FL88) is typed and mailed to the contract monitor requester. The information is then entered into the AQ data base, and copies of the letters are filed.

One of four events should then occur:

- The document arrives at DTIC and starts through the pipeline. When descriptive cataloging is complete, a clerk from the Bibliographic Data Base Branch searches the AQ data base and, if the document is on order, notifies the Acquisition Section. The Acquisitions Section then updates the AQ data base and hardcopy file and notifies the requester of the AD number.

- A denial note is received from the contract monitor. In that case, the Acquisition Section updates the AQ data base and hardcopy files and notifies the requester.

- No notice is received. After 90 days, the Acquisition Section sends a follow-up letter on the basis of information from the AQ data base. The AQ data base and hardcopy files are updated with this information. After another 90 days, a "no response" is treated as a denial and the requester is notified.

- The source notifies the Acquisition Section that the document is not yet available. The Acquisition Section notes this information in the files and makes another request when the source said the document would be available. The requester is notified of the revised availability date.

Each of these paths should (within 180 days) result in a closing of the AQ data base record with either an AD number for the acquired document or a denial (the exception being a document that will become available later).

The operation of the Acquisition Section is one of the methods by which DTIC acquires documents. In general, DTIC relies on receiving documents from the source as part of the primary distribution. However, the actions of the Acquisition Section are separate from those of the main TR processing operation. At present the data captured in the AQ data base are not shared with RTIS, the system that supports on-line entry of TR citations. The only interaction between the Acquisition Section and the sections in the processing pipeline is the notice from Bibliographic Data Base Branch reporting receipt of the document. Document inputting begins with arrival of a document in the mail room.

2.3.4 Mail Room (DTIC-FPS)

The mail room receives all incoming DTIC mail. Classified documents are signed for, every document is stamped with the date and a sequential number, and placed in a cart for delivery to the Selection Section. All other mail is delivered to individual offices.

2.3.5 Selection Section (DTIC-FDAC).

The Selection Section reviews newly received documents and evaluates their suitability for inclusion in the TR data base. Documents received from the mail room are transferred from the mail cart onto holding shelves to await selection. During this transfer, if a document already has an AD number on the cover or seems especially old, it is set aside. Any accompanying correspondence is reviewed, and the document may be duplicate-checked. This is done to quickly isolate duplicates or documents DTIC has requested from external sources to refilm due to deteriorating microfilm at DTIC. If the documents fall into either of these categories it is removed from the input processing flow.

Staff shortages and increased acquisition efforts have resulted in a backlog of documents on the holding shelves. The backlog now amounts to nearly 600 documents, representing approximately 5 days of pipeline production. Documents are taken from the holding shelves by selectors, generally on a first-in, first-out basis. They review the documents for technical content. If a document is unacceptable, it is returned to the source. Fewer than 1 percent of the documents received are rejected.

Selectors also review every page for defects in physical quality, such as unreadable copy and missing pages. Distribution statements are also checked carefully for correctness. Distribution statements and other information used throughout the cataloging process are included in the "Report Documentation Page," (DD Form 1473). These distribution statements are generally prepared by the report author(s) and should accompany every document sent to DTIC.

The selector completes the review by attaching instructions for further processing. Examples of these instructions are: direction to return borrowed copies of documents or to film compendium articles as separate documents. All acceptable documents are then reviewed briefly by the senior selector and placed on a table for forwarding to the Bibliographic Data Base Branch. As part of their work, selectors maintain logs of their production. These logs are collected and sent to management daily.

As a result of the physical review, approximately 15 percent of the incoming documents are found to have problems that prevent immediate cataloging. These documents are routed to another staff member of the Selection Section, who resolves these problems by calling or writing the source. Up to 45 days is allowed for resolution of problems. A document whose problems are not resolved in that time is returned to the source. Approximately 95 percent of the problems are satisfactorily resolved, and the documents are returned to the pipeline.

In addition to its regular duties, the Selection Section is also responsible for several special functions. One is to review new input into the data base by remote SBIN users to identify documents that may not have been submitted to DTIC for secondary distribution. Every Technical Abstract Bulletin (TAB) cycle (2 weeks), the Selection Section receives a list of all documents input by SBIN users. This list is reviewed by Selection Section staff members, and any recommendations for acquiring documents are sent to the Acquisition Section to be ordered. The list is held by the Selection Section. When an SBIN document is later received, the Bibliographic Data Base Branch requests that page in the list for review in preparation of an entry in the data base.

The Selection Section also receives requests from outside sources for file maintenance. These are forwarded to the Data Base Support Branch for action.

In addition, the Selection Section also processes replacement documents. Documents may be replaced at the request of the source because they have been superseded or contain outdated information. These documents are identified either by the accompanying correspondence or by the duplicate check performed by the Bibliographic Data Base Branch. The original AD number, if unknown, is obtained through a DROLS inquiry and, together with the new document, is sent to microfilming to replace the original. In addition, a "Change in Status of Document" (DTIC Form 28) is completed and sent to the Data Base Support Branch for file maintenance of the TR data base.

The Selection Section also processes requests from the Micrographics Division to re-obtain documents that were originally recorded on microfilm and are now to be placed on microfiche. The Selection Section contacts the source and, when the document arrives, routes it back to the Micrographics Division.

2.3.6 Bibliographic Data Base (DTIC-HDB).

The Bibliographic Data Base Branch is responsible for creating the computer citation containing the descriptive bibliographic data from the document. After initial processing by the Selection Section, a clerk in the Bibliographic Data Base Branch counts the documents and records information about document categories (e.g., classified, unannounced). This count is recorded on an "Incoming/Outgoing Documents" form (DTIC Form 444) and verified with the Selection Section. The documents are then placed in appropriately marked envelopes. The envelope markings denote the classification level of the documents and the TAB cycle they are part of. The clerk then places the documents in another holding area, where a senior cataloger sorts the documents into separate stacks of approximately equal cataloging work.

As catalogers become available to do new work, they select stacks to process. They first search the TR data base for each document, to find out whether it is already in the collection. A second search of the "current file" is also performed. The current file contains all references to documents in the pipeline (not yet updated into the master TR data base). The current file is updated twice a day for duplicate-checking purposes and, near the end of the TAB cycle, may contain as many as 2,000 citations. It is purged every 2 weeks with the DROLS update and grows again in the next TAB cycle.

The duplicate check of the TR data base and the current file should result in one of three conditions: finding no duplicate citation, finding a matching citation submitted by a SBIN site, or finding a duplicate. If the document is a duplicate, it is marked with its AD number and sent to inventory, thus completing the process.

If the document being cataloged is an SBIN document, it is turned over to an SBIN specialist in the Bibliographic Data Base Branch for review. The specialist obtains a copy of the SBIN listing from the Selection Section, marks any corrections or modifications needed for conformity with DTIC cataloging standards on the listing, and assigns a new AD number and enters the changes on the terminal. The document then goes to another staff member, who performs the subject indexing by recording additional subject terms on the "Document Data Worksheet" (DTIC Form 41). SBIN documents are processed in the Bibliographic Data Base Branch; non-SBIN documents are sent to the Subject Analysis Branch for subject indexing. The document and DTIC Form 41 are then forwarded to the Subject Analysis Branch,where the information is entered via cathode-ray tube (CRT) into the RTIS. The Selection Section is notified of the change in AD number so that a supersession transaction can be initiated.

More commonly, a document is neither a duplicate nor an SBIN input. It is therefore processed or "worked-up" through the rest of the cataloging and input steps. This process includes identifying and formatting the information to be transferred from the document into the data base. Such items as title, personal author, corporate author, security classification, and page count are extracted from the document and entered into the bibliographic citation via RTIS. Subject-oriented data are excluded from this processing step.

As part of this process, the source header book or card file is checked for the correct computer code for the corporate source. The rules that govern descriptive cataloging are documented in the "DTIC Cataloging Guidelines," which is 90 pages long and defines the entry of some 30 fields.

Many of the data input formats are complex. For example, if a document is authored by Milton B. Smith and John D. Jones, their names are entered into RTIS as follows:

M. B. /Smith;J. D. /Jones

Once the work-up is complete, the cataloger keys the data into the RTIS. Specific methods of working up and keying in a document vary among catalogers. Some complete most of DTIC Form 41 and key from it; others fill out only a small portion of the form and key the rest of the data directly from the document. Keying data into RTIS requires use of a CRT to enter a series of commands, field number tags, and the actual data. The system provides no menus, prompts, or formatted screens. The user is required to know and type all commands (e.g., @NI@ is the command to initiate a new document) and all field tags (e.g., @10@ is the personal author field tag), as well as the data formats. An example of how a typical document citation would appear is shown in Figure 2-6.

When the entry of a document is complete, the @SI@ (store item) command is entered. The system then displays edit-checks, security field contents, field lengths, and character types. The operator reviews the results of these checks and may either continue to edit the document or proceed to another. A printout of each entry is given to another cataloger for review.

Once document citations are entered into RTIS, a clerk in the Bibliographic Data Base Branch searches for them in the AQ data base. If a match is found, a printout is sent to the Acquisition Section to update the AQ data base and to notify interested users that the document is now available from DTIC. Documents are then

FIGURE 2-6. SAMPLE CITATION INPUT VIA RTIS

```
@1@ADA156701
@1A@A
@2@p8/4
@6@Arguments against Alleged Proof of the NA-K Pump
in Studies of K + and NA + Distributions in Amphibian Eggs,
@10@K. L. /Gibson;W. P. /Boggs
@11@1983
@12@13
@15@N00014-78-C-0125,$PHS-2-R01-CA46341-03
@20@u
@21@Pub. in Physiological Chemistry and Physics NMR, v18 p314-325 1984.
@23@*EGGS, *CATIONS, *SODIUM, *POTASSIUM, *AMPHIBIANS,
TRANSPORT PROPERTIES, AMPHIBIOUS OPERATIONS, BLOOD PLATELETS,
CELLS (BIOLOGY), DISTRIBUTION, EQUATIONS, ERYTHROCYTES,
FROGS, HUMANS, HYPOTHESES, LIFE(BIOLOGY), MUSCLES, NERVE
FIBERS, PARTICLES, PUMPS, SOLUTES, THEORY, MEMBRANES(BIOLOGY),
TRANSPORTS PROPERTIES, REPRINTS.
@27@With the aid of an ingenious technique called the reference phase method,
White and his coworkers tested the alternative theories of solute distribution in
living cells. Many interesting and significant findings led us to the belief that their
claim of a proof for the Na pump theory was unwarranted. By taking into account
other relevant data, one can argue that their data, in fact, support an alternative
theory, association induction. Distribution patterns of K +, Na + and other solutes
in all kinds of living cells, including frog muscle, human red cells, and amphibian
eggs, as well as all subcellular particles, including yolk platelets, can be adequately
described by one general equation of solute distribution.
@28@u
@29@2
@30@Reprint: Arguments against Alleged Proof of
the NA-K Pump in Studies of K + and NA + Distributions in
Amphibian Eggs.
@33@1, 20
@35@400229


@end@
```

Notes: This is a complete citation, including subject indexing. The weighted terms are preceded by an asterisk. This format, which starts each field on a new line, is one method of inputting and displaying RTIS data. Alternatively, the citation is simply entered in a continous stream:
@1@ADA156701@1A@A@2@p8/4@6/@Arguments against...@end@

counted, and the count is recorded on DTIC Form 444. The documents are sent to the Subject Analysis Branch and Form 444 is submitted to management for inclusion in the daily pipeline report.

In the process of working up documents, catalogers must consult the Corporate Source Header List or a manual card catalog to determine which computer code represents the corporate source of the document.

Occasionally the source of a document is not in the list. When this happens, the cataloger forwards the document to a specialist, who establishes a new source header and corresponding computer code. To establish the new header, the specialist first researches reference materials to identify the source, and then completes a "Corporate Author Update" (DTIC Form 242). This identifies the new source, along with its source code, sort code, and geopolitical code. Information on Forms 242 is keypunched every 2 days, and the header file is updated on the computer. The computer generates a transaction listing for review by the specialist who created the source header and code. A 3x5 card is also prepared and filed. On average, 15-20 new or modified headers are added every 2 days. Every 6 months, a complete alphabetical list of source headers is printed for DTIC use. A similar list is also generated for external users once every 5 years (annual supplements are also provided).

Once the new source header has been established, it is placed in the hierarchy of organizations by another specialist. When information about the source is received, various reference tools are checked to determine the hierarchical position of the organization. After that position is determined, the specialist updates the hierarchy by using a listing that prints out all sources in hierarchical order. That listing displays the source header codes and their corresponding hierarchy codes but does not list the source name.

The specialist then fills out an "Electric Accounting Machine (EAM) Transcript Transmittal" (DSC Form 24) with the commands to locate, change, or delete the header. The forms are keypunched by the Telecommunications and ADP Systems Directorate (DTIC-Z) and reviewed by the specialist. If the cards are correct, the update is run, and a new listing is printed for review by the specialist. The process is repeated until the hierarchy is accurate. Because of the complexity of the updating process, several iterations are often required. At long intervals, the system also produces a lengthy alphabetic print-out of the hierarchy.

2.3.7 Subject Indexing – Subject Analysis (DTIC-HAS).

The Subject Analysis Branch (DTIC-HAS) is responsible for entering the data that will provide subject matter access to the citation. Access to TR data base citations is provided through the title, the document abstract, and three types of index terms.

The first type of index term is the Committee on Scientific and Technical Information (COSATI) fields and groups, a subject-indexing scheme that assigns codes to major (fields) and secondary (groups) areas of knowledge. Although not directly searchable, COSATI fields and groups indicate the general subject of the document.

The second type of index term is the DRIT, a set of controlled terms developed by DTIC that more closely reflects DoD terminology and interests.

The third type consists of open-ended, uncontrolled terms. Typically, these terms are either specialized ones (words or phrases) selected by the indexer and used in some narrow field of research or they are specific names or acronyms, such as B-1 (bomber). The proper selection of indexing terms is important for information retrieval access because the abstract is not directly searchable by the DROLS retrieval system.

The COSATI fields and groups and the open-ended terms are selected manually by the indexer. Selection of DRIT terms, however, is assisted by MAI. MAI consists of a series of sophisticated computer programs that accept input text (abstract, title, etc.) and parse it into selected words and phrases that are then mapped onto the most appropriate terms in the DRIT. The MAI process has been in use for a number of years with the WUIS and IR&D data bases but has only recently been applied to the TR data base. The MAI process is now the center of operations at the Information Analysis Branch.

Documents received in the Subject Analysis Branch from the Bibliographic Data Base Branch are first counted and sorted. SBIN documents that are already indexed are sent to an analyst for review of subject codes and then to keying. Compendiums and documents in microfiche/film format are assigned to an indexer, who performs the work manually without help from MAI. All other documents are assigned to indexers on the basis of subject matter.

The indexer reviews the document abstract and reduces it to the maximum allowed length or, if it is too short, expands it. Where possible, the author-supplied subject terms are appended to the abstract. The indexer also notes on the DTIC Form 41 any translation notes. The abstract is then given to a data transcriber, who enters it into the RTIS.

The MAI processes the data overnight without reviewing or editing the input for quality. When the print-outs are received the next day, they are broken apart and where possible, given to the indexer who first marked up the abstract.

For classified documents, the MAI displays the recommended terms on the printout but does not post them in the computer record. The indexer reviews the listing and indicates which terms are to be deleted and which are to be added. Any weighted (primary) terms are also noted on the listing. During this work, the indexer has access to the document to assist in review. The listings are then given to the section supervisor for review.

For unclassified documents, the MAI generates the same printout but also posts the terms in the citation on the computer. In this case, the indexer relies entirely on the listing, adding and deleting terms as needed. Unclassified documents are not reviewed by the supervisor.

For both classified and unclassified documents, the indexer also records COSATI fields and groups and uncontrolled terms on the MAI sheet. The marked-up MAI listing and the document are returned to the data transcriber for entry/ modification of the subject terms into RTIS. Entry is a complex process. For classified documents, the transcriber must enter the terms with any weighted terms first. For unclassified documents, the situation is more complex because the computer record already contains the MAI-posted terms. Those terms must now be rearranged, with the weighted terms placed first and some of the other terms deleted or added. The transcriber must either perform a complex edit or key in all the terms again. At that time, fields and groups, translation notes, and uncontrolled subject terms are also entered. The document is then flagged for edit processing. Once entered, the documents are counted and noted on an "Information Analysis Branch Daily Activity Report" (DTIC Form 82) for use in the pipeline report. They are then sent to the Data Base Support Branch.

2.3.8 Data Verification (DTIC-HDS).

The Data Base Support Branch verifies the accuracy of the data entered by the Bibliographic Data Base and Subject Analysis Branches. In addition, the Data Base Support Branch is responsible for releasing citations to the actual TAB processing and DROLS updates.

Every night, edit programs process the document citations completed by the Subject Analysis Branch and produce listings for delivery to the Data Base Support

Branch the next morning. The seven basic printouts are broken down by document type (unannounced, SBIN, etc.). Of the seven, the three primary listings are the Print-E's, Print-R's, and the Text Verifier. The Print-E is a complete display of every document citation and includes messages to highlight errors detected by edit program. The Print-R is the same as the Print-E for document citations that have no detected errors. The Text Verifier uses a portion of the MAI process to detect spelling errors in the title, abstract, and other selected fields.

When the listings are received, selected information is copied from the summary listings (e.g., missing AD number listing, the list of invalid subject terms) marked onto the Print-E and Print-R lists. These lists are then broken apart and matched with the documents or with the preceding day's printouts of recorrections. A proofreader then carefully compares the listing with the document and the DTIC Form 41 and DD Form 1473 and notes any errors on the listings.

Next, the Text Verifier is consulted and the text is proofread; again, any errors are marked on the Print-E/R. A document is reviewed only once. It is then sent to microfilming; any additional proofreading consists of comparison with the marked-up listings. If no errors were detected in any of the fields used to produce the microfiche header (author, title, etc.), the AD number is listed on a "Daily Microfiche Header Sheet" (DTIC Form 386). At the end of the day, these forms are collected and entered into RTIS to trigger the running of a program that processes the headers for microfilming.

If errors are found in a document citation, they are corrected by data transcribers on the RTIS, and a flag is set to rerun those items through the overnight edit programs. The listings are stored for comparison with the results of the edit-checks the next day. When a document becomes error-free, that fact is noted in the RTIS by the absence of a flag, and its listings are stored away. Listings are held for approximately three TAB cycles and then destroyed.

Other statistical reports that are produced daily include the "count-out", a report listing the processed citation AD numbers by category (e.g., classified, unannounced), and the "TAPERS" report, which provides counts of the number of transactions (additions, changes, deletions) processed against the TR data bases.

Another daily statistical report is the "overlay" report, which is particularly important because it detects the loss of any documents. Like any computer system, the RTIS requires a unique identifier for every record in its files. For RTIS, this is called the "document name," and DTIC uses the AD number as that name. Whenever a new item is created, the AD number is entered with the @SI@ command, thus naming the document. If, however, there is already an item in the current file with that number, RTIS replaces the existing stem – without warning – with the new document data. Since AD numbers are unique, duplicates occur only when one of the numbers is entered incorrectly. Since the AD number is also included within the document citation, the overlay report prints out any document citation in the Current File if the document name and the internal AD number disagree. This listing helps detect many overlay errors, but the problem still must be investigated and the overlayed document citation must be reprocessed.

The activities of the Data Base Support Branch are a reflection of the document processing cycle leading up to publication of the TABs. Each TAB consists of citations for documents processed during the preceding 2 weeks. To keep pace with the volume of incoming documents, the production pipeline at DTIC must catalog an average of 120 documents a day.

Because the Data Base Support Branch is at the end of the pipeline, it receives few documents early in the cycle, but the numbers increase as the cycle progresses. Every day, it completes some citations and must recorrect others. As the cycle nears the end, the Branch has accumulated a large number of correct citations as well as a growing number of errors that must be corrected. In the very last days of the cycle, the workload may be heavy and it must be completed under the pressure of the TAB deadline. Once all the document citations in a given TAB are verified, the Data Base Support Branch informs the Operations Branch (DTIC-ZOA) that the Operations Branch can begin the TAB publishing and master AD file (MAD) updates. Once the updates are completed and the citations are available in the TR file on-line, staff members from the Data Base Support Branch use RTIS to delete each of the roughly 1,200 citations in that TAB from the Current File. This procedure completes processing of new TR documents by the Data Base Services Directorate. However, the directorate must also process later changes in documents.

2.3.8.1 File Maintenance.

Processing of later changes in documents, referred to as "file maintenance operations," is controlled by two staff members in the Data Base Support Branch. They process three major types of changes: security classification upgrades or downgrades, superseded documents, and miscellaneous changes. Miscellaneous changes occur when an error – typically a misspelling or typographical error – is noted in a citation. When the File-Maintenance Group is notified about an error, the item is extracted from the TR data base and corrected in RTIS. The next day, a Print-E or -R is received and verified. Once corrected, the revised item is processed as part of the current TAB for the TR data base update but does not appear in any of the announcement products.

Supersessions usually occur either when DTIC is replacing the SBIN AD number with a DTIC AD number or when the source is replacing a document with a new version. (Changes in security guidelines are a major example of this type of replacement.) Supersessions are usually processed by the Selection Section (DTIC-FDAC), which forwards a completed Form 28 to the R&E Programs Data Base Branch. That branch then reduces the old AD citation in the TR data base to a cross-reference to the new document, which will have been processed in the normal flow.

2.3.8.2 Security Information Processing.

Information related to citation and document classification and distribution is maintained in more than eight different fields in the TR data base record, and similar information is kept in a separate data base on the Sperry 1100/82 computer. This latter data base is the inventory file, which contains the order processing and document inventory records. It can identify every user who orders a given document, a capability of particular importance for processing upgrades in document classification.

Document or citation classification upgrades are initiated upon receipt of a written request from the source agency. A standard file-maintenance transaction is used to update the DROLS data base. This updating includes not only changing classification levels, but also recording the TAB cycle during which the change was made (in the open-ended terms field) and adding the reason for the change in Field 49. While some security-related data are automatically copied to the inventory data base from the TR update, other information is updated directly by completion of an "Inventory Change Card Sheet" (DTIC Form 25) and keypunching of the transactions. A classification upgrade also requires production of a printout listing from the request-processing file identifying users who have received the document. This list is produced by completing a "Management Information Inquiry System Worksheet" (DTIC Form 381) from which computer cards are keypunched. All document holders on the list are notified by letter of the classification upgrade.

A report classification can be downgraded in three ways: at the request of the sponsor, automatically by date, or automatically by event. Automatic downgrade by event, however, has not been used to date by DTIC. Requester downgrades occur in exactly the same fashion as upgrades except that downgrading changes are announced in TAB rather than by letter. Date-related downgrades are identified by periodic printing of the "Automatic Downgrade" report, a list of all documents that have passed the downgrade date entered in the citation when the document is first received by DTIC. Those documents are processed through file maintenance to change their classification in DROLS. Any change in the security fields is detected by the edit programs and causes two actions. The first action is the previously discussed movement of a copy of that information to the inventory file where a "before-and-after" change report is generated to verify the change. The second is generation of an "ATNCR24" report to account for the changes in the files used to accumulate new or changed citations awaiting updating into the TR data base.

Security changes are highlighted in a section of the TAB publication. A rough copy of this section is generated in the "S20040" report, which displays the exact changes as they will appear in the TAB. The result is proofread by the file-maintenance group. These TAB sections are also accumulated and published in an annual summary of security changes. This step completes security processing, except that DTIC also provides requesting libraries with a yearly summary of security changes for documents cataloged in the TR data base.

2.3.9 Shared Bibliographic Input Network Processing.

The SBIN enables remote DoD technical libraries and information centers to enter citations directly into the TR data base by means of the RTIS. SBIN sites may use the same terminals for document input as they use for the search of DTIC data bases.

Access to DTIC takes one of three forms: classified access over dedicated communications lines, unclassified access over dedicated lines, and unclassified access over dial-up communications lines. More than 50 technical libraries and IACs are now capable of participating in SBIN. In 1984, those sites accounted for 10 percent of the total citations entering the TR data base. (WUIS data base inputs were also made by some SBIN sites.) To assist these organizations and DoD in general, the SBIN project is intended to accomplish the following:

- Issue early announcements of new reports to the DoD community
- Provide more detailed and accurate subject indexing by local experts
- Share document processing costs within the DoD community
- Enable each SBIN site to use the TR data base as its local catalog by adding its local holding symbol to the record.

This last feature is a powerful capability for libraries that do not maintain their own automated catalog. To use the feature, the site enters a complete citation plus the holding symbol assigned to its local library. SBIN sites need not enter full citations. If they check the TR data base or the Current File for duplicate citations and find the citations already in the DTIC data base, they may simply add their holding symbols to the record. They may also perform this check by submitting keypunch cards or magnetic tapes to DTIC. Some sites prefer this latter method of loading large numbers of holding symbols to reduce the amount of time spent on-line.

The IACs have the same capability through the use of fields set aside specifically for IACs. A separate subject field is included exclusively for IAC use. The SBIN also provides the sites with DTIC printouts that can serve as hardcopy catalogs of their collection.

The holding-symbol capability is essential for libraries that do not have automated, on-line catalogs, but none of the other features offer any significant direct advantages to a site performing input. For libraries that maintain their own systems, SBIN input represents a duplication of cataloging effort. Therefore, many SBIN sites do not input document citations, preferring to mail their documents to DTIC and let it do the cataloging.

Where sites do perform SBIN inputs, they perform them in exactly the same manner as internal DTIC input. Items are cataloged in accordance with DTIC cataloging guidelines (which frequently differ from local practice). RTIS is also used as in DTIC; however, since SBIN members use RTIS far less than DTIC catalogers, the lack of formatted screens and other tools tends to lengthen the time required to enter a citation. When citations are entered, they are transmitted to DTIC and enter the same RTIS files as do citations for internal pipeline documents. They become part of whatever TAB cycle is currently in process.

The SBIN document citations are processed overnight along with the other citations, and the edit lists are printed the next morning. Members of the Data Base Support Branch review and mark the listings as time allows. They then separate the lists by contributing site and mail the lists back to the sites. An SBIN site reviews the list and may make corrections in a file-maintenance transaction. Because SBIN citations become part of the TAB cycle, they are updated into the TR data base as part of that cycle. Depending on when the SBIN document citation is entered, it may be updated the following night (perhaps with errors) or held up from entering the TR data base for as long as 2 weeks.

Even SBIN sites that enter citations themselves must still send the documents to DTIC to be filmed and made available for ordering. These documents become part of the DTIC pipeline. They should be recognized as SBIN documents when a descriptive cataloger duplicate-checks them. As described in Section 2.3.6, SBIN cataloging is reviewed and changed as necessary. DTIC also reviews SBIN-entered citations for the potential addition of subject terms and changes the AD number to show that the document is available from DTIC. This work is further reviewed by the Data Base Support Branch as a part of the pipeline process. Hence, SBIN citations still require significant DTIC effort before being loaded into the TR data base.

2.3.10 Research and Engineering Program Data Base (DTIC-HDR).

WUIS contains data on research and development projects being conducted by the Military Services. The IR&D data base includes data from independent projects run by DoD contractors (industry, academia, etc.). Both data bases are managed and controlled within DTIC by the R&E Programs Data Base Branch in the Data Base Services Directorate. That Branch has minimal interaction with the other organizations doing this work in DTIC. The WUIS and IR&D inputs are also processed in a substantially different manner from TR data base inputs. Data are still entered through RTIS, but all the batch programs and manual processes involved in the input process are different. Moreover, the output products are different, and there is no production cycle corresponding to the TAB.

2.3.10.1 Work Unit Information System Input Processing.

The WUIS data base consists of approximately 180,000 records, of which some 26,000 contain data on active projects. In a year, approximately 10,000 new projects are initiated and an average of 70,000 projects undergo changes and close-outs. WUIS inputs arrive at DTIC by several different media, including magnetic tape, punched cards, on-line entry, and hardcopy "Research and Technology Work Unit Summary" (DD Form 1498). Of the total input, 95 percent reaches DTIC in the form of magnetic tape.

Tapes are formatted as card-image transactions, as defined in DoD Manual 3200.12-M-1, "Research and Technology Work Unit Information System," published by DTIC. Despite the standards defined in the manual, the tapes arrive in a variety of physically different forms [e.g., some use American Standard Code for Information Interchange (ASCII), and others use the Extended Binary Coded Decimal Interchange Code (EBCDIC) format]. DTIC must therefore maintain several different programs to read these tapes. These conversion programs must be modified whenever changes in submitter's hardware, software, or personnel cause tape format changes. The quality of these tapes also varies; tapes from some users are regularly returned as unreadable. Tapes are submitted periodically by the Services; the distribution of work is fairly even through the year.

When a tape is received at DTIC, a clerk logs it in and determines which conversion program to run. The tape is then submitted to computer operations for conversion to Sperry-Univac format and printing of a list. This list is reviewed by the Management Information Branch staff to ensure that it is in the correct format and can be processed by the update program. Limited amounts of WUIS input also come from on-line input (from external sources, much like SBIN citation entry) and direct submission of Form 1498 in hardcopy. On-line data input also results in generation of lists. Both on-line and tape listings are mailed to the submitter for editing and correction. Hardcopy Forms 1498 are marked by DTIC control clerks and editors and submitted to data transcribers for entry into RTIS. The mark-up consists of both reformatting the data into DTIC format and adding the field tags, as for TR inputs.

All inputs for the week are submitted for update on Friday night. The update program accepts all transactions and applies them against the old master file tape and generates a new master tape. The update program also generates the "CSL Listing," several sets of statistics, and an inverted file-transaction tape. All new transactions on the new master tape are read by the MAI program, which posts the terms into the RTIS file. The CSL list is a formatted print-out of each item for which a transaction occurred (i.e., new record, change to existing record, or deletion of existing record).

In addition to the transaction record, the CSL prints messages for any edit/audit errors detected by the update program. The update program breaks down these errors into major and minor errors. Records with minor errors are processed into DROLS; records with major errors are not. Both types of errors are noted in the

CSL list. Staff members from the R&E Programs Data Base Branch review the CSL and statistical reports on Monday. If everything is acceptable, they direct computer operations to load the new inverted and master files to DROLS that night.

During the week, any errors indicated on the CSL for inputs from the R&E Programs Data Base Branch are corrected, but errors from on-line or tape submissions are not corrected. Instead, these erroneous listings are mailed back to the submitters for correction. Thus, some of the data loaded into the data base may contain errors. Even the errors corrected by the this Branch are not added to the data base until the next weekly update.

The output listings from the MAI process are also printed over the weekend and forwarded to the Subject Analysis Branch for review. Because the Subject Analysis Branch does not have enough staff members to review more than a few items, most are posted with the terms assigned by the machine. MAI postings in RTIS are also processed in the following week's update, completing the processing cycle for WUIS inputs.

2.3.10.2 Independent Research and Development Input Processing.

The IR&D data base consists of more than 71,800 records. An average of 8,000 transactions takes place each year, but unlike the situation with WUIS, the work is not evenly distributed over the year. More than 60 percent of the IR&D reports are received between April and August, corresponding to the budget cycle followed by the 80 submitting organizations. These organizations have some 400 suborganizations that contribute reports. Submissions are all in the form of hardcopy "Independent Research and Development Data Sheets" (DTIC Form 271). Some organizations are considering switching to tape.

IR&D input is processed in essentially the same manner as WUIS hardcopy input. Submissions are received and logged in by a proofreader in the Management Information Branch. They are then forwarded to a control clerk and editor for marking up. This mark-up consists of adding accession numbers and RTIS field indicators and converting data into DTIC formats. Data transcribers then enter the data into RTIS. These data are also updated every Friday night by the same programs as WUIS but with different parameters and tape files. For IR&D data, the CSL is proofread by an R&E Programs Data Base Branch proofreader and corrections made in RTIS for the next week's update. IR&D data are also processed by MAI programs with minimal manual intervention.

2.3.11 Management Information.

In any production environment, adequate management information is critical for identifying backlogs for which additional staff or overtime may be required or areas with excess resources that can be used elsewhere. This information must be available quickly to reflect the current status and must be presented in a form that can also be used to establish long-term trends.

For TR processing, management information is displayed in the "Daily Activity Report - Document Input Processing" (DTIC Form 102), more commonly referred to as the "Pipeline Report." The report, released daily by the Policy, Plans and Resource Management Office (DTIC), is compiled from production records and logs maintained by individual staff members involved in the input process. Supervisors aggregate data for their organizations, combine them with data from available computer-generated reports, and submit the compilation daily to the Policy, Plans and Resource Management Office. These reports are submitted in several different forms. The Bibliographic Data Base Branch, for example, submits DTIC Form 373; the Subject Analysis Branch uses DTIC Form 82. The Policy, Plans and Resource Management Office aggregates reports manually from all of DTIC and produces the daily Pipeline Report. Once the daily report is developed, it is distributed to DTIC managers.

2.3.12 Data Processing (DTIC-Z).

The Telecommunications and ADP Systems Directorate (DTIC-Z) is responsible for all data processing in DTIC, providing both programming and operations support for the production data bases (TR, WUIS, and IR&D). All data processing is now run on two Sperry computers. The Sperry 1100/82 is the primary computer, processing DROLS searching, RTIS, and the batch input programs for all three data bases. For the functions discussed in this FD, the smaller Sperry 1100/61 processes the AQ data base and the ordering inventory data bases. The 1100/82 is seven years old, and since its installation, processing demands have increased dramatically, necessitating a number of upgrades. The system is now being upgraded for additional disk storage. The CRT terminals used by RTIS are as old as the computer and frequently fail because of overheating. Replacement terminals are being purchased.

Despite its upgrades, the 1100/82 system still reflects the state of technology in 1978 at the time of its installation; the situation is illustrated by the need to back up some of the mass storage devices to tape every half-hour. System response time is criticized by both inside and outside users. DTIC plans to replace the system with new equipment through competitive acquisition, but the system cannot be replaced until all the Sperry-unique software is rewritten or converted to machine-independent American National Standards Institute (ANSI) Common Business-Oriented Language (COBOL). This conversion process is referred to as "software normalization."

DTIC's computer capacity will be enlarged next year with installation of Gould 9050 computers to support the Defense Logistics Agency (DLA) (DMINS) office automation project. The ability of these computers to support such DTIC operations as the input system has not yet been established.

DTIC software includes more than 550 programs, some 300 of which are in Sperry assembler language. The RTIS was developed 11 years ago; DROLS and much of the supporting software are even older. With the exception of RTIS input and DROLS retrieval, most of the programs are batch jobs, running overnight and using tape files rather than disk. Documentation for many programs either is outdated or does not exist. Furthermore, personnel turnover within the Telecommunications and ADP Systems Directorate has led to a loss of experience with the undocumented programs.

The age of the hardware and software slows the operations of the input system. At present, almost all intermediate files used in processing are kept on magnetic tape. Most data base update and edit programs run overnight and produce tape or list outputs (the lists are themselves tapes, which are printed). The tapes are then processed by the next job program in a long sequence of programs required to complete updating of the TR data base. This environment leads to a complicated tape-handling effort, especially when the highest tape density is 1600 bits per inch (bpi) (the TR master file, for example, now occupies more than 40 reels of magnetic tape).

The software normalization effort, which is now in its first stages, will address many of these problems. Normalization of the TR data base input programs will not only convert code from assembler language to COBOL but will also convert many of the tape files into disk files. The effort will also produce programs written in a more generalized manner and using modern coding techniques, current documentation, and some user-requested modifications. Although completion of the effort will have a beneficial effect on computer operations, it will not affect a number of important areas. The RTIS itself will not be changed; the supporting batch programs, although streamlined and operating with disk files, will remain overnight-batch programs; and WUIS, IR&D, and all the auxiliary files will remain unaffected.

2.3.12.1 Technical Report Processing.

The processing of TR data base computer records by the input system is shown by flow charts in Appendix C. This section provides a summary description of the data processing steps involved in updating the TR data base.

All data entry is done through RTIS, an on-line interactive program that enables the user to input and modify document citations. It can also retrieve document citations from the DROLS master TR data base for modification. As noted before, RTIS is entirely command-driven. It provides no menus and few on-line edit-checks or other assistance.

Document citations entered into the RTIS are maintained on two disk files. The first contains the text of each citation and the second is a directory of document citations. For purposes of this discussion, it is important to note that the directory contains processing flags that are input by the terminal operators and indicate to the batch programs what processing is to occur for a given document citation. For example, when a document citation is to go through edit-check processing by the Technical Reports Branch, a "/B" flag is keyed in; when a document is to be deleted from RTIS, a "/D" flag is entered.

Every night, report citation records held in RTIS are read by the first of some 20 programs involved in their processing. The next two programs divide these citations into as many as seven different files, including current TAB, previous TAB, IAC, and unannounced. These files are then processed separately by the succeeding job steps, which include the edit-check programs, text verification, MAI processing, and inverted current file update. As the last step of each night's processing, each of the seven files is merged into the previous night's tapes of document citations in process. These files, referred to as the "MINI-MAD" files, continue to accumulate documents until the end of a TAB cycle. At that time, each final MINI-MAD is used to generate its output products and is then merged into the master TR data base.

Little information is available on the processing of auxiliary files, such as source acronyms, headers, and hierarchy files. The study team did not obtain detailed breakout of the MAI programs, but the DTIC Long Range ADP and Telecommunications Plan (AD-B083 896) lists more than 20 COBOL program modules. The AQ data base file is maintained on the BASIS data base system on the 1100/61. No detailed list of these programs was obtained, but discussions with staff members from the Systems Design and Development Division (DTIC-ZD) indicate that they consist of a data base definition and a few simple listing programs. The AQ data base file does not interact with any other file. The DRIT is in the early stages of being put on-line, by means of the UNIDAS document retrieval system.

2.3.12.2 Work Unit Information System and Independent Research and Development Processing.

WUIS and IR&D inputs are processed by the same programs and are treated here as a single process. Figure D-6 in Appendix D depicts the process flow described in this section. Each incoming WUIS tape is run through one of several tape conversions to create a Sperry Univac format. RTIS data are also converted into a format acceptable to the update program. All transaction tapes are merged and records sorted into accession number sequence. The merged and sorted transaction tape is combined with the old master file tape to create an updated master file tape. A tape of new/change/delete inverted file transactions is created, along with the CSL print-out and a separate set of transaction statistics. These reports are sent to the R&E Programs Data Base Branch. The new master file tapes (now 11 reels for WUIS) are processed through the MAI programs. MAI processes only new items found on the tape, along with those specified by parameter card entries (although it must read all 180,000 WUIS items). The MAI list is created and sent to the Information Analysis Branch. The terms generated by the MAI are also posted to the citation files in RTIS.

When the R&E Programs Data Base Branch approves the update, the inverted file transactions are sorted and run against the old inverted master file, and a new one is generated along with a statistical report. The new master file and new inverted file are then loaded onto disk and made available to DROLS. Any output processing is not covered in this FD.

2.3.13 Summary.

The present environment for creating a TR data base citation is extremely manpower-intensive. A typical document is handled by five organizations in two directorates. It changes hands at least 153 times and often as many as 203 times. Parts of the citation are keyed into a terminal during at least eight different data entry steps. On average, a document requires more than 3 weeks for processing and adding its citation to the TR data base.

Through SBIN, document citations can be transmitted electronically to DTIC, but their edit listings are returned to the user by mail. Because of this delay, citations are frequently processed into DROLS with errors that can not be corrected until the SBIN user has received the edit listing and keyed in the changes. Once a copy of this document arrives at DTIC, it is renumbered—and sometimes altered to conform to DTIC data formats—and undergoes additional subject indexing. Although the SBIN program is intended to supplement DTIC cataloging resources, participation is a burden on the SBIN site and still requires work by the DTIC cataloging staff.

The processes for operating and updating the WUIS and IR&D data bases are simpler, but they also rely on the same data processing technologies as the TR input, that is, predominately batch processing and manual editing. Feedback to SBIN users also relies on printouts mailed to the originating site. It requires at least 2 weeks to process a document completely into DROLS.

2.4 Proposed Methods and Procedures.

The replacement input system will serve all the DTIC operational functions supported by the current input and AQ data base software, and will extend support to some functions that are currently performed manually. For each functional area, it will provide improved data-processing support and reduce the demands now placed on the DTIC staff. Details of the proposed flow will be provided in the system specification.

2.4.1 Summary of Improvements.

The following list summarizes the more significant improvements in the input operation that can be expected from the new system:

- Fewer manhours will be expended to acquire a new document
- Better acquisition reporting will be provided to assist increasing acquisitions
- Fewer manhours will be expended to input a document
- Accuracy of citation data will be improved
- Quality of indexing will be improved
- Less elapsed time will be required to complete input processing
- Fewer manhours will be spent processing duplicates
- SBIN participation will be increased
- External input of documents through additional media will be increased
- Data lost through failure of hardware/software or through human errors of users or computer operators will be reduced
- Personnel will be better utilized by having the routine tasks
- Flexibility to meet future growth and changes in hardware, software, applications, and input volume will be increased
- Availability of programming support and productivity for system enhancements and new development will be increased

- Fewer manhours will be spent on management information reporting and better management data will be available to track pipeline status and productivity.

2.4.2 Summary of Impacts.

2.4.2.1 User Organizational Impacts.

The system will have a minimal impact on the organization. The proposed work flow recommends that the duplicate-check process be transferred from the Bibliographic Data Base Branch to the Selection Section. It also suggests that the Data Base Support Branch functions change from 100 percent review to sampling techniques; that change would alter the make-up of the organization from a large number of clerical personnel to a smaller number of higher graded personnel. Recently, the Data Base Support Branch has acquired the data entry personnel from the Subject Analysis and R&E Programs Data Base Branches. The proposed system can accommodate this function in either a distributed or centralized mode. The proposed work flow will establish one new function—that of document completion. Based on the current organization, this work could be performed in the Data Base Support Branch.

2.4.2.2 User Operational Impacts.

The business functions to be performed by the user organizations will remain the same. However, the actual processing techniques will be greatly improved, resulting in a better product at a lower cost. Since the amount of time expended per document processes will be reduced, the requirements for personnel time will also be reduced. That savings will be critical if DTIC is successful in increasing its rate of acquisitions. Personnel savings can also be used in areas that currently need more resources, such as indexing WUIS and IR&D data.

2.4.2.3 User Development Impacts.

The proposed input system is intended to replace a currently operational system. During the development stage the user will retain the full capabilities of the current system. The new system will be implemented by means of a series of parallel processing phases as each data base is implemented on the new system.

Potential impacts on the user during actual programming or system implementation will depend on the decision on the hardware approach. If the new system is placed on the current mainframe on which the operational system is now resident, some disruption of service may occur. If it is placed on a newly acquired stand-alone minicomputer, the new machine can be completely dedicated to development until the first application is placed into test production.

SECTION 3. DETAILED CHARACTERISTICS

3.1 Requirements Summary.

The summary of requirements presented here for the replacement input system is derived from the proposed methods and procedures recommended for data input by the DTIC staff, SBIN members, IACs, and other remote contributors. During the next stage of system development, alternatives for implementing the requirements will be formulated and evaluated and a method of implementation will be recommended. Therefore, the summary of requirements and accompanying proposed methods and procedures will form the set of system characteristics used for subsequent system design and development. Requirements for the replacement input system are divided among four purposes here: improvements in user access and productivity, management information and control, system operations, and software support and development. In addition, accuracy and validity requirements and timing constraints are presented. These are summarized in the following subsections.

3.1.1 Improvements in User Access and Productivity.

- User access via terminals offering menus, formatted screens, function keys, and other techniques that reduce the need for memorizing system commands. These capabilities should be provided for both new input and file maintenance.

- On-line edit-checks of entered data and immediate feedback of the results to the user. Edits should include checks for length, format, acceptable values, user-defined edit routines, required data, and cross-field checks. These checks should also be available for batch (e.g., tape) loads.

- Spelling check capabilities which operate in either on-line or background mode with results displayed on-line. Spelling check must operate on large fields of text, and must easily accommodate new words.

- On-line availability to update, display, and verify input data against auxiliary (authority) files. These input data include the source header, source hierarchy, source acronym, DRIT, and fields and groups.

- On-line duplicate-checking of the TR data base, in-process citations, and AQ data base citations in a single query. Matches found in the AQ data base file should automatically produce a file update, acknowledging receipt of the document and notifying interested patrons.

- General on-line query access to files maintained, created, or updated by the input system applications. Search access should include all fields of the files. Users should be able to enter searches easily.

- Improvements in MAI processing, including:

  - Spelling verification before processing

  - Sorting and elimination of duplicate terms before and after final review of the terms

  - Simplified method of applying and editing terms in final review

  - On-line verification of posted terms

  - Automatic feedback to monitor MAI performance

  - Automatic movement of weight terms to front position.

- Consolidated verification reports. The number of reports should be modified and all aspects of the document should be edited. The reports should display the data in convenient, easy-to-read formats. Whenever possible, edit listings should be exception reports so that users can edit and list document citations on the basis of optional selection criteria.

- Control of user access and security down to the data field level within a computer record. Access authority should include the ability to specify read-only, read-write, or prohibition of access. It should also include selected authority to permit original entry of data and subsequent updates or edits.

- Control of assignment of document numbers and prevention of overlays of citations.

- Automatic production of form letters throughout document acquisition, selection, and reclassification.

- Automatic production on demand of standard reports and regular production reports.

- Transfer of bibliographic records from the AQ data base to the replacement system for further data entry during cataloging.

- Formatted data entry screens for use throughout the cataloging process. Full-screen edit features for entering and correcting citations and other entries.

- Acceptance of on-line, volume input of citations created by external sources and transmitted to DTIC. Edit and verification results for these citations should be retained on-line for review and correction by the organization submitting them.

- Global updating of IAC subject fields to reflect changes in subject vocabularies.

3.1.2 Management Information and Control.

- Capture of production information by monitoring of system inputs or by alternative technologies, such as bar-code tracking of documents.

- Extensive report-generating capabilities on production data. Production data to be selected, sorted, and displayed as needed. Displays to include tabular and graphic formats. Reports to include status and long-term trends. Standard reports to be run automatically.

- Earliest possible capture of basic bibliographic information to provide current, accurate information about the status of the process.

- Information regarding the place of any document in the pipeline.

3.1.3 System Operations.

- Support for 80 to 180 DTIC staff and remote users at a time. Support for a mix of letter-quality and high-speed printers with upper-and lower-case capabilities.

- System availability of 98 percent from 6:00 a.m. to 8:00 p.m., Eastern Standard Time (EST).

- On-line storage capacity to meet requirements of current data bases and provide for future growth.

- Telecommunications capability for a variety of terminals, printers, and other computers.

- Software capable of residing on Sperry computers or on hardware capable of interfacing with them.

- Acceptance of a variety of input media, including CRT, OCR, diskette, tape, and batch transmittals from microcomputers.

- Ability to write to resident disk storage and tape drives, as well as printers and CRTs.

- Capability of tape drives for variable recording densities, including 1,600 and 6,250 bpi. The system need not accept keypunch data, and all internal applications should avoid such input.

- Satisfaction of DoD and National Security Agency (NSA) requirements for maintaining classified data up to the secret level.

- Minimal operator intervention and use of tape files in standard production processing. The system should be capable of storing and running extensive sequences of multiple-program jobs.

- File backups and other protective measures to minimize the effect of user or operator errors and hardware malfunctions.

3.1.4 Software Support and Development.

- Structure for growth in:
  - Number of applications
  - Size of current applications
  - Size and number of data fields in current applications
  - Number of users.

- Use of a data base management approach to:
  - Simplify the enhancement and modification of current applications
  - Simplify the development of new applications
  - Standardize programming techniques across applications
  - Minimize programming support for the operating system and concentrate on applications instead.

- Complete current documentation along with vendor training and support.

- Structure that minimizes machine-dependency.

3.1.5 Accuracy and Validity.

The DTIC input system will primarily be receiving, retrieving and displaying bibliographic data. Maintaining the validity of these data will primarily be a function of machine edit-checks of the data at the time of input. Numerous types of edits checks will be required, including but not limited to the following:

- Minimum and maximum length

- Required or not required

- Numeric, alphabetic, character position by position (masking)

- Table and authority file look-ups including an extensive spelling check capability

- Valid date formats
- Validate the content of one field based upon the contents of another
- Availability of exits to user-defined edit routines.

These edit-checks will be applied on-line to data being entered onto a CRT and in batch mode for data transmitted by tape or other off-line format.

Numeric computations which will mostly be included in management reports should be accurate to a level of two decimal places for use in percentile representation.

3.1.6 Timing.

The timing constraints placed on the input system are the ability to:

- Respond in 15 seconds or less to an average on-line interactive query of three data bases with up to 50 users working at any given time
- Respond in 5 seconds to an average command
- Complete the following functions:
  - On-line edit-checks of data within 20 seconds
  - Complete spelling checks of 300-word abstracts in a background mode (5 minutes), or interactively on-line
  - Complete MAI of document abstracts overnight or in background (30 minutes)
- Produce standard production and management reports overnight
- Produce ad hoc reports in less than 4 hours
- Produce tape backups of the disk files every 4 hours
- Respond to data base schema modifications on an overnight basis
- Restore the data base or its associate software to a different baseline overnight
- Recover from an intermittent system failure within 2 hours
- Recover from a catastrophic system failure overnight.

3.2 System Functions.

This section discusses the individual system functions as they relate to the user functions described in Section 2.4 and the specific performance requirements in Section 3.1.

3.2.1 Data Input.

The primary system function will be to input and store data that will later be used by processes outside the scope of the system. The bulk of the data input will be through CRT entry at DTIC, with the remainder coming in a variety of formats (tape, OCR, etc.) both at DTIC and remote locations.

3.2.1.1 On-line Input.

Most data will be entered by being transferred from a hardcopy document to the system via a CRT. Two types of screens are to be used for this process. One will be a formatted screen with field names fixed on the display and blank spaces for the user to enter the data. The second will be a blank screen where the user streams data and field separation characters. In either mode, the system will perform substantial data validation of the entered material. These edits shall be performed in real-time and their results shall be displayed on the screen. Errors shall be indicated and the user given the facilities to correct them.

In addition to the field-oriented data described above, lengthy document titles and abstracts will also be entered.

This system must provide convenient text-handling facilities and an extensive spell-check feature.

Some of the work will also include on-line editing of previously entered data. Approximately 10-20 percent of the TR data will be entered through CRT terminals in remote locations. Those entries are to be given the same features as internally entered data.

3.2.1.2 Batch Data Entry.

Currently more than 90 percent of the WUIS data and limited amounts of the IR&D data is received on magnetic tapes. The system needs to submit these tapes to the same edit processes as for on-line input and make the results available in a file for on-line correction. When documents are displayed on the screen, detected errors will be indicated.

At this time, no other batch inputs occur, but the system should be capable of accepting and editing the following types of input:

- Microcomputer floppy diskettes
- OCR equipment output
- Electronic transmission in batch mode from remote computers
- Bar code reader wands, combined with CRT input.

3.2.2 Query Processing.

Users will perform Boolean type searches on the contents of the fields in the system. The system must support an inverted file or some other structure suitable for rapid retrieval. The system should support searching on any combination of fields to the word level. Retrieved documents should be displayable on the CRT or routed to a printer. The user should have control over the fields, formats, and order of the displayed data.

Two special types of query processing are required for this system. The first is to search multiple data bases with a single search and have the results combined for display on the users terminal. The second special process is to include amongst those queries, queries of the main TRs data base on DTIC's Sperry-Univac 1100/82 computer.

3.2.3 Reports.

The system should produce a series of standard reports. Some of these reports should be generated automatically based upon parameter-driven schedules. Such reports would be generated on a daily or weekly schedule usually overnight. The remainder will be requested as needed through user terminals.

The system must also provide an ad hoc report-generating facility where non-data-processing users can easily develop and run production reports on an as-needed or regular basis.

3.2.4 Data Transfers.

The DTIC input system will be used to prepare information for loading into data bases to be accessed by the DROLS retrieval system and to generate a variety of information products from currently existing software. The input system will therefore need a capability to convert data from its internal format into other specified formats. Once the conversion is complete the data can be transferred to DTIC's Sperry-Univac 1100/82 computer via magnetic tape or batch electronic transfer.

On a periodic basis document citations will be extracted from these data bases and downloaded to the input system for editing and revision. These downloads will be requested by document identification number (AD number) entered through user terminals. The request will have to be transferred to the data base, and the document will have to be extracted, converted to input system format, and made available to the user to edit.

Query processing, as described in Sections 2.4 and 3.2.2 will also occur across multiple data bases, including data bases controlled by the input system and those under control of DROLS. Users will formulate the query and indicate the data bases to be searched. The system should then automatically process the query against the selected data bases and return the results in a single query file. This process is to be performed in a real-time mode.

The input system and the DROLS data bases both access large authority files such as the DRIT and the Source Header files. These files will be maintained and updated on the Input Systems, and must be periodically copied to the DROLS system. These transfers can be performed on a nightly basis whenever the files are updated.

3.2.5 Security and Administrative Functions.

The DTIC input system will contain document citations classified up to the secret level. Since the system will be accessed by users with and without clearances, security requirements will be critical.

The system should operate at the B-2 level of NSA guidelines for a trusted computer.

The system must provide security down to the field level. Privileges to access documents and fields should include no access, read only, and read-write access as follows: update only, create only, and full access. These privileges should be authorized based on combinations of user ID, terminal ID, organization code, and user password as compared to the security tables and the classification level of the citation.

3.2.6 Data Base Functions.

The basic component of the system should be a generalized data base management system and should include all features commonly associated with such a system. These features include but are not limited to:

- Supporting multiple data bases
- Creating, updating, and deleting records, data bases, and application programs and supporting files and tables
- Providing for data backup and failure contingencies.

3.3 Inputs/Outputs.

These sections present brief descriptions of the inputs, outputs, and data used to perform the functions of the DTIC input system. This section lists the input/ output (I/O) and what functions it supports.

3.3.1 Inputs.

The primary means of data input to the system will be through CRT screens formatted for specific functions. In actual development more than one physical screen may be required for each function. In all cases, screen functions include data validation, validation feed-back, and editing features. The screens needed for the input system include:

- Acquisitions screen – input data for requesting a document. This screen will include bibliographic data identifying the document and the document source. Completing this screen will automatically generate a document request form letter. This same screen will be used later to display and edit the request status.

- Duplicate-check screen – will be used to initiate duplicate-check queries. Depending on the results of the query, the data input will serve as the basis for a new input system record.

- Selection problem screen – will input data on documents that cannot clear the selection process. They are placed on hold until the source is contacted and the problem resolved. This screen includes basic bibliographic data about the document and the source as well as an indication of the problem. This screen will also generate a letter to the source.

- Bibliographic cataloging screen – will process the basic bibliographic data for the document. This screen will also generate a second duplicate-check query. It will operate in two modes: a formatted fixed screen mode and a streamed text mode.

- Document abstract screen – will process the document abstract in a text editing environment.

- Post review screen – will display and provide editing capability on MAI terms. It will also allow other data to be entered.

- Quality control/ file maintenance screen – will display and provide editing for all fields of the record.

- Authority file screen – will provide for the maintenance of the various authority files in the system.

- WUIS input-edit screen – will input and edit all WUIS data and text fields.
- IR&D input-edit screen – will input and edit all IR&D data and text fields.
- Generalized query and report screens – will screen and make menus available to utilize the ad hoc query and report generation features.

The system will accept batch inputs that will be equivalent to the fields entered for the WUIS and IR&D screens, and the TR bibliographic, abstract, and post review screens.

3.3.2 Outputs.

The system will typically provide three types of outputs: reports (standard and ad hoc), file transfers, and screen displays.

3.3.2.1 Screen Displays.

Most of the screen displays are reactions to screen inputs (such as queries), and are covered in Section 3.3.1. This section discusses only those screen displays that are initiated by a batch program. They include:

- MAI post review – This screen is also an input screen. It initially displays the DRIT terms suggested by the MAI program.
- Batch inputs – The result of the edit validations of batch inputs to the WUIS, IR&D, and TR data bases will be placed in a file that can be edited by the appropriate input screen.

3.3.2.2 Batch Reports.

The following standard batch reports will be included:

- Form letters – For document requests, request follow-ups, and problem resolutions.
- Proof lists – A well-formatted report showing all fields of a document citation and highlighting any detected errors. One proof list will be needed for each data base and one for the authority files. The TR data base proof list can be used by all branches.
- MAI report – This report will be used by the vocabulary control personnel to evaluate and modify the MAI program performance. It includes the document abstract, the machine chosen terms, and the final indexer selected terms.

- AD number report—Listings of all AD numbers assigned by the system and the corresponding input system item number; a label is also provided for affixing to the document.

- Microfiche header list—Lists by item number all citations included on the microfiche header tape.

- Citation completion list—A nightly report of AD numbers completed and released for update with a separate section of items that have not been completed and have passed the allotted time for completion.

- Management activity report—A daily status of the pipeline activity.

- Long term trending report—A periodic report of activity.

- Group/individual productivity report—A periodic report of productivity.

3.3.2.3 File Transfers.

This section includes all regular manipulation and transfer of data between the input system and other DTIC systems and functions.

- Inverted file transactions - The primary purpose of the input system is to provide complete and accurate citations to be loaded into the appropriate DROLS data base (TR, WUIS, etc.). On a regular basis—currently 2 weeks for TR and weekly for the others—citations will be converted from input format to DROLS update format and transferred to programs within that process.

- Authority file updates - The authority files will be updated and maintained on the input system; however, they will also be accessed by the DROLS system and other related software. Therefore, these files will be copied to that system at the end of any day in which an update occurred.

- Microfiche header tape - A tape generated on a daily basis contains microfiche header data for all documents successfully completed that day.

3.4 Data Characteristics.

The function of the input system is to provide cost-effective and accurate bibliographic citations for later retrieval by other systems. The data are primarily bibliographic rather than numerical and should seldom be altered once correctly entered.

3.4.1 Data Base Growth.

The TR, WUIS, IR&D, and problem data bases are transactional data bases that will grow to a certain point and then be transferred to the permanent data base while the input data base is purged. For example, the TR file will grow at an approximate rate of 120 citations a day for 10 days. At that time these documents will be uploaded to the TR data base under DROLS and erased from the input system.

The AQ data base will be a permanent file and will grow steadily. It currently contains 14,000 records and will grow at a rate of 200 records a month.

The authority files will be the largest files in terms of number of records; however, these files will be primarily static, with only the source header and hierarchy file showing any growth.

The system must be able to support the convenient adding of additional data bases to meet future DTIC requirements.

3.5 Failure Contingencies.

An important aspect in the development of the input system is the ability of its resources (personnel, hardware, and software) to respond to failures. The requirements for responding to failures or potential failures are described in the ensuing subsections.

3.5.1 Backup.

To minimize the amount of time lost from damage to, or errors in, the data base and its associated software (e.g., program libraries, directories, tables), sets of back-up files will be created and maintained. These files will be created using installation standards.

3.5.1.1 Data Base Backup.

An image of the data base will be produced at the time that the data base is declared operational, and it will be modified periodically. This image will be duplicated on both disk and tape.

How often the data base will be backed up (copied), by whom, how many versions will be kept, where the backup versions will be archived, when the data base will be purged, and documentation (audit trail) of the contents of each version will be determined by the system manager or Data Base Administrator (DBA) at a later stage of the system development. However, backups of the primary data file will be at no more than 4 hour intervals. All backups of the data base will be created with password protection to ensure that the proper access with read-only permission is granted.

The backup software and its job control language should be kept as a cataloged procedure to ensure data base integrity and security.

3.5.1.2 Data Base Management Software Backup.

An image of the software (source statements, executable modules) and its associated job control language used to process data, produce standard reports, and assist the users will be created when the software is declared operational, when software enhancements are made, and on a periodic basis. All failure contingencies discussed concerning data base backup apply to the DBMS software.

3.5.2 Restart.

The DBMS selected to operate and maintain the DTIC Input System data will be able to restart execution of a batch update at the latest checkpoint once the program terminates because of a program or system failure. The frequency of checkpoints will depend upon the hardware and software environment.

The data-base-restore capability will be executed when the data processing staff determines that a program or system failure can be classified as an intermittent failure. In an intermittent failure, the data base contents remain intact and the data base is readable. The data processing staff will use the version of the data base that was being used at the time of the failure and reapply all transactions to the data base. This action will be initiated as soon as possible after error detection in order to restore the use of the system to all input system users.

SECTION 4. DESIGN DETAILS

4.1 System Description.

The replacement input system is designed to improve the system functionality for the user, improve performance reliability, and reduce the ADP level of effort to maintain and enhance the system. The system being replaced is composed of a single interactive program and a large number of batch programs all written in assembler language more than 10 years ago. As proposed, the replacement system will make use of a commercially available DBMS that has the basic system requirements already inherent in it. It should also contain system development tools to simplify the creation of DTIC's data bases, screens, and reports and forms, and to perform other system-specific functions. Details of the operation of the proposed software approach are provided in Section 3, while Section 5 provides some alternative hardware approaches.

4.2 System Functions.

The DBMS software is expected to provide inherent functionality in three major areas:

- Data base functions, including file structure, disk storage techniques, updating methodology, security, failure protection, and other features common to such systems.
- Development tools such as a program development language for the easy development of reports, forms, and queries. These tools will include methods for the development and implementation of data bases and supporting files. Development capability will include provision to interface with specially developed features such as linking to DTIC's MAI programs and extracting documents from the Sperry computer.
- User tools, including a report generator and interactive query language oriented to non-data-processing users. The structure of the data supporting the use of the query language shall be of an inverted file or equivalent technique to permit rapid searching of large amounts of data through Boolean parameters.

System documentation, training, and installation, maintenance and application support should be available from the supplier. Details on timing, accuracy, and validity are provided in Section 3.

4.3 Flexibility.

At this time, DTIC should plan to incorporate four major data bases within the input system: the AQ data base, TR transaction, WUIS transaction, and IR&D transaction data bases. Additionally a number of supporting files should be included, such as source headers and hierarchy, fields and groups, MAI, and others. The system must have the flexibility to permit the easy incorporation of changes to these data bases. Changes can include adding or subtracting data elements, changing data length and formats, and modifying screens, queries, and formats. Changes will also occur in the interfacing of the system to other DTIC systems and production run procedures. The system should also permit the addition of many other data bases.

4.4 System Data.

Data inputs and outputs are described in Section 3.

SECTION 5. ENVIRONMENT

5.1 Equipment Environment.

This section defines the hardware and software required to support system modifications required for the implementation of an improved input system. Detail provided in this section are given at a more detailed level than is normally provided in an FD. This additional detail is provided for two basic reasons.

First, the DTIC input system described in this FD is intended to replace a currently existing multicomponent system that has been operating for more than 10 years. The replacement system must be interfaced with current hardware and software that will remain in place and work in conjunction with the input system.

Second, two hardware approaches are being considered. One is to place the system on DTIC's current mainframe, and the other is to place it on a stand-alone minicomputer. Because of the complexity of interfacing with the current system, we performed a detailed evaluation of the viability, desirability, and costs of both approaches at this stage of the project development.

The minicomputer evaluation was performed by selecting two specific machines that exhibit a range of characteristics, a Digital Equipment Corporation (DEC) VAX II/750 and a Sperry Model 11. Research was conducted on both systems to determine how they would interface and communicate with the current mainframe. These machines were selected only to provide a "proof of concept" and not to establish recommendations for specific vendors. The mainframe approach was evaluated on the assumption that an unspecified commercially available DBMS would be placed on the Sperry 1100/82 coresident with other DTIC systems.

As a part of this evaluation, in which DTIC data processing staff members participated, numerous factors were considered, including such issues as whether the current DTIC mainframe computer could support an additional central processing unit (CPU). This section describes the current environment to document the resolution of such issues.

A description of the proposed configuration for all three systems – the two minicomputers and the mainframe – is provided in this section. Additionally, details on the hardware and software environment of the remaining DTIC systems to be interfaced with is also provided. This detail is included to serve as background data for the selection and to serve as a technical basis for implementing the selected alternative.

The central computers that DTIC now operates are Sperry models 1100/82 and 1100/61. The DROLS function is executed on the 1100/82, and the 1100/61 is used to manage a variety of applications including the TR AQ data base. As currently configured, these mainframes do not provide the entire range of processing and support functions required for efficient operation of the input system. This limitation is due in part to the batch processing techniques used to support its operation.

The improved system design described in this section will retain many of the characteristics of the current system; however, at a minimum, it must satisfy the following requirements:

- Available 98 percent of the time from 6 a.m. to 8 p.m., EST.
- Support for 100 simultaneous users with the capacity to grow to 180 simultaneous users within 3 years.

Approximately 66 percent of users will access the system over dedicated lines, and the remainder will be dial access.

Since the proposed input system replacement program places a strong emphasis on a need for effective real-time, on-line data processing and improved user access to files, computer system hardware and software modifications will be necessary.

Three system design alternatives are presented in this FD. The first alternative is based on the continued use of the existing Sperry 1100/82 hardware and includes replacement of the existing RTIS software with new input system software. The second alternative is based on the addition of a Sperry System 11 minicomputer, and the third alternative is based on the addition of a DEC VAX 750 minicomputer. All three alternatives require the development of new input system software offering the functions described earlier in this report.

The three alternatives selected include the entire range of possible actions that could be taken for the enhancement of the input system, from the use of the existing system hardware with enhanced software through the additional processor power to provide the required levels of responsiveness, reliability, and expandability for the new system. Two minicomputer alternatives were examined to identify the range of costs and options associated with the implementation of the new system. In both cases, a detailed design and cost estimate was developed to provide DTIC with the assurance of the feasibility of acquiring a usable minicomputer system from more than one vendor.

In all cases, the new input system software includes the use of general-purpose DBMS software to support the file management activities of the input system, and in all cases, the AQ data base will be removed from the Sperry 1100/61 and added to the functions of the new on-line input system.

The software system design for all options includes similar security processing features. These features have been developed in a manner that duplicates the security features of the existing system.

5.1.1 Alternative 1: Sperry 1100/82.

This subsection provides a detailed description of the existing system hardware. This hardware configuration will be used to support the operation of the enhanced input system software, defined as Alternative 1. The 1100/82 consists of two CPUs, main storage units (MSUs), storage interface units (SIUs), system transition unit, system maintenance unit, I/O units, console and peripherals.

The CPU performs the arithmetic, logical, and instruction sequencing operations and interfaces with the system transition unit, maintenance unit, I/O units, and storage interface units.

The I/O units control all transfer of data between the peripheral devices and main storage and consist of two sections—a control section and a channel section containing I/O channel modules. The DTIC 1100/82 has two I/O units and two I/O expansion units. Each I/O unit has 16 word channels plus expansion units with a byte multiplexer channel and block multiplexer channel.

The control section includes all logic circuitry associated with the transfer of function, data, and status words between main storage and the subsystems.

Main Storage: The main storage system consists of MSUs and SIUs. The 1100/82 is configured with four random access memory MSUs, each of which has an expansion module. The MSUs and the two expansion units each provide 262K words of memory. This system arrangement provides the 1100/82 with a total memory capacity of approximately 2.1 million words. The SIU consists of moderate capacity, high-speed storage buffers. With its expansion units, it provides 16K words of buffer storage.

System Transition and Maintenance Units: The system transition unit contains the controls and indicators for partitioning the central control complex units and operator panels for each of the CPUs and I/O units. The system maintenance unit provides diagnostic routines and fault isolation for the CPU and the I/O units through automatic comparison of the internal logic status against known correct data. The system maintenance unit includes a maintenance processor, card tester, communications capability, and a CRT work station.

Auxiliary Storage and Peripheral Subsystems: Auxiliary and peripheral subsystems are attached to I/O channels. The 1100/82 offers three media for data storage: cache/disk, magnetic disk, and magnetic tape.

The Sperry cache/disk system is a mass storage subsystem that, for most applications, provides the host computer system with extremely fast response to mass storage I/O requests. The cache/disk system uses both disk and solid state storage. The system configuration includes a 5057 Cache/Disk Processor, a 7053 Storage Unit, and Disk Units. These units are attached to the 1100/82 through a word channel interface.

The 1100/82 system provides word channel interfaces to the 8433 Magnetic Disk Subsystem. The 1100/82 has six 5046 Control Units providing access to three 8433 Subsystems containing various sets of disk units. The sets contain 8, 13 and 14 disk storage units, respectively, providing a total storage capacity of 7 gigabytes (Gbytes).

The 1100/82 system uses a byte channel as an interface with the magnetic tape subsystem. This subsystem includes two 5042 Control Units and 14 Uniservo 36 Magnetic Tape Units. The Uniservo 36 tape unit provides a group code recording scheme with a recording density of 1,600 bpi at a tape speed of 200 inches per second (ips).

Input Devices: Data are entered into the system using magnetic tape, key punch, or remote terminals. The terminals used for system input include:

- On-line/Dedicated Terminals: DROLS has 155 terminals directly connected (dedicated) to the 1100/82 system. These terminals operate in a synchronous mode at 9,600 bits per second (bps).

- Dial-up Terminals: DROLS has 732 synchronous and asynchronous terminals connected to the 1100/82 system on a dial-up basis. Synchronous terminals operate at 2,400 bps and asynchronous terminals have multiple rates.

Output Devices: System outputs are displayed on printers and CRTs. The output devices include:

- On-line/Dedicated Printers: The 1100/82 uses a Perry 0776 Printer subsystem as an output device. It is a free-standing, medium-speed, self-contained unit.

- Dedicated Terminal Outputs: 155 CRT terminals are directly connected (dedicated) to the 1100/82 system. These terminals operate in a synchronous mode at 9,600 bps. An 800 series matrix printer is used to make hardcopies of message data sent to the CRT.

- Remote Outputs: 1100/82 users can download or prepare data information at Uniscope 200 terminals supported by Sperry 610 cassette tape decks.

Distributed Communications Processor (DCP): The existing Sperry equipment includes a DCP/40 communications processor. It is described here because it plays a significant role in all of the system alternatives.

Although primarily designed for Sperry products, the DCP/40 system offers the flexibility to interface with a broad range of communications and computer equipment provided by other vendors. Its architecture uses multiple processors designed to function as a distributed processing system. It includes a general-purpose communications processor for network control and general message management.

Up to 16 specialized I/O processors can be used for line and terminal handling. They present a standard programming interface for controlling data transfer over serial lines, parallel channels, or host processor connections. Processor programs carry the same software design for all types of equipment. The microprocessor-based system can be adapted to the communications environment of other hardware systems.

A variety of line modules is available to meet a number of communication protocol and speed requirements. Representative capabilities include asynchronous to 9,600 bps, synchronous to 19.2 K bps, and wideband to 64 K bps. The synchronous protocols that can be accommodated include Universal Data Link Control (UDLC), International Business Machines Corporation (IBM) BISYNC, and Systems Network Architecture Synchronous Data Link Control (SNA/SDLC). Line modules will interface with the following standards; for asynchronous and synchronous communications, Electronic Industries Association (EIA) RS-232C, V.24, and V.28; for high speed communications, V.35.

The DCP/40 offers two line modules for host interface, a byte-channel module and word-channel module. The word-channel connection operates in a dual-channel mode combining 16-bit parallel connections to form a 32-bit information transfer medium. The byte-multiplexer channel connection offers both a single and dual-channel version. Both interfaces can only be connected to Sperry equipment. The host interface modules provide speeds of 1.2 Megabytes per second (Mbps) for both the byte-multiplexer channel and the word channel.

The DCP/40 operated by DTIC includes the following features:

- 165 ports.
- Line speeds: The DCP/40 currently communicates with Uniscope 200 polled synchronous terminals at transmission speeds of 9,600 Kbps for in-house users and 2,400 Kbps for dial-up users. It also supports asynchronous users at multiple transmission speeds.
- Protocols: The DCP/40 system as currently configured provides an EIA RS-232/CCITT V.24 interface with asynchronous and synchronous protocols.

<u>Communications Equipment</u>: Communication equipment supporting the DROLS system includes modulating/demodulating (modem) equipment, telecommunications security equipment, and land telephone lines.

- Modem Equipment: Various types of modems are used for all terminals external to the Defense Documentation Center (DDC). Modems provide an interface between telephone lines and the computer system. Both dial-up and dedicated telephone facilities are utilized. All ports operate at a 2,400 bps rate of speed. System network design includes some multiplex circuits.

- Telecommunications Security Equipment: A large number of system terminal users are authorized to receive classified data. Therefore, data encryption is utilized for security integrity. KG-13 and KG-34 cryptographic equipment with the associated Crypto Ancillary Unit are used for this purpose. The DDC also provides a crypto-out-of-synch detector.

5.1.2 Alternative 2: Sperry System 11.

The use of a minicomputer will shift a portion of the Sperry 1100/82 mainframe workload to the minicomputer. File data input and updates will be predominantly accomplished on the minicomputer. However, file inquiry traffic will occur between the minicomputer and the mainframe computer. The minicomputer will make inquiries and transfer file data from the mainframe to its files.

Periodic total file updates will occur between the files of the minicomputer and the mainframe. The update period is file-dependent and will vary from daily to biweekly. Table 5-1 shows the files that will be transferred periodically from the minicomputer to the mainframe.

TABLE 5-1. FILES THAT WILL BE TRANSFERRED FROM MINICOMPUTER TO MAINFRAME COMPUTER

FILE	SIZE	PERIOD
Source Header	198 Kbytes	Daily
Source Hierarchy	5.1 Mbytes	Daily
Fields & Groups	12.6 Kbytes	Biweekly
DRIT	87.4 Mbytes	Biweekly
Lexical Dictionary	34.6 Mbytes	Daily
RecognitionDictionary	10.2 Mbytes	Daily
TR Transaction	30 Kbytes	Biweekly
WUIS Transaction	10 Kbytes	Weekly
IR&D Transaction	10 Kbytes	Weekly

One of the two minicomputer alternatives considered during the development of this FD was the Sperry System 11 computer. This equipment was considered because of its compatibility with the existing Sperry equipment. The System 11 is capable of executing all of the existing system software, and can be readily interfaced with the existing DCP/40 processor. Thus, the use of this equipment will assure a high level of equipment compatibility and minimize the difficulties associated with a multi-vendor environment.

The remainder of this section describes the Sperry System 11 minicomputer. Again, while specific system characteristics are presented in this section, these characteristics are intended only to be representative of the type of equipment that is capable of supporting the input system requirements. The use of the System 11 and the DEC VAX minicomputers is not intended as an endorsement of those particular types of equipment.

Central Processor Unit: The System 11 central processing equipment is housed in a single cabinet. A general-purpose instruction processor executes all programs. Its features include a 128-word general register set, 32 base registers, and an instruction set that is completely compatible with the Sperry 1100/82. The System 11 minicomputer processor system supports the same operating system as the 1100/82 which uses Sperry 1100 operating system.

The System 11 will support channel I/O processors, which provide a byte-channel, block-multiplexer-channel and disk-controller-channel interface. The system can provide a block-multiplexer channel level link to the Sperry DCP/40 front-end processor.

Main Storage Unit: System 11 main storage consists of control logic and memory chips housed in a basic cabinet. The recommended configuration is equipped with 1M word of main storage. Main storage memory is expandable in 1/2M word increments to accommodate future growth.

Auxiliary Storage and Peripheral Subsystems: To support the anticipated file storage requirements for the minicomputer system, a minimum of 165 Mbytes of on-line peripheral storage will be required. To provide for additional data base expansion and the storage of operating system, utilities, and applications software, the initial system should be configured with at least 300 Mbytes of auxiliary on-line storage capacity. The following peripheral storage units should be included in this configuration:

- Disk Units: To support immediate and future growth requirements, we recommend that the System 11 (or equivalent) be equipped with an 8,436 disk subsystem (or equivalent). The subsystem should be configured with one controller and two disk units. This configuration provides 460 Mbytes of storage capacity.

- Tape Units: The system offers one 2014-99 data streaming tape unit for making copies of large quantities of data to support off-line storage and manual transfer of files. The 2014 is a 9 track, 1/2 inch tape, 1,600 bpi, phase encoded unit, with a transfer rate of 160 Kbps at 00 ips.

Input/Output Devices: System I/O can be accomplished utilizing existing terminals and printers.

System 11 Interfaces: File inquiries and file transfers will occur between the System 11 and 1100/82 as a function of normal operational file inputs and updates. These transfers will be made using the interface between the System 11 and the DCP/40. As has been previously noted, files will be transferred periodically from the System 11 minicomputer to the 1100/82 mainframe as batch transfers of data. This file transfer activity will necessitate a communications link with a high throughput capability.

As set forth in Section 2, the level of file inquiries and updates is expected to be as shown in Table 5-2.

Based on the file transfer schedule, an aggregate of approximately 150 Mbytes of file data will have to be transferred electronically at least every 2 weeks. Since the data transfer rate between the System 11 and the DCP/40 is a very high

TABLE 5-2. EXPECTED LEVEL OF FILE INQUIRIES AND UPDATES

FILE	ESTIMATED INQUIRIES/UPDATES PER DAY
Acquisition	350
Source Header	200
Source Hierarchy	10
Fields & Groups	600
DRIT	600
Lexical Dictionary	200
Recognition Dictionary	200
TR Transaction	1260
WUIS Transaction	200
IR&D Transaction	150

1.2 Mbps, the total time required to complete this file transfer will be 125 seconds, and it can be done anytime after the end of operational hours.

The DCP/40 provides a host-to-host interface capability through its line modules. The 1100/82 mainframe will interface with the System 11 using a block multiplexer channel between the System 11 and the DCP/40, through the DCP/40 front-end processor to a word-channel connection with the 1100/82. This interface will support a data transfer rate of 1.2 Mbytes. The same interface will be used to support communications with data terminals connected to the DCP/40.

Telecommunications Considerations: The interface between data terminals and either DROLS or the new input system will be provided in the same manner as it is today. Terminal users will have to alter sign-on procedures to select the system (DROLS or the input system) that they wish to access.

Distributed Communications Processor: No modifications are required for the DCP/40 with the exception of the addition of a block-multiplexer channel.

5.1.3 Alternative 3: DEC VAX 11/750.

The second minicomputer alternative evaluated during the preparation of this FD was the DEC VAX 11/750. (This minicomputer is referred to in the remainder of this description as the VAX 750.) The VAX 750 was selected because it offers commonality with other DTIC equipment and because of the wide variety of available DBMS software. However, other minicomputers also provide similar functions.

The functional relationship between the input system, operating on the VAX 750, and the remaining DROLS software operating on the Sperry 1100/82 will be similar to the relationship described for the System 11. File management and updates will be predominantly performed on the VAX, with periodic file transfers between the VAX and the *mainframe*. In addition, the VAX will process terminal input traffic and will make periodic inquiries to data bases residing on the 1100/82.

The following text describes the VAX 750 minicomputer system architecture.

Central Processing Unit: The VAX 750 CPU is a 32-bit, synchronous microprogrammed computer that executes variable-length instructions in native mode and nonprivileged procedure definition processor (PDP)-11 instructions in (PDP-11) compatibility mode. The processor's features include 32-bit virtual addresses allowing access to more than 4 Gbytes of virtual address space and 16 (sixteen) 32-bit registers.

The internal bus structure is used for communication between the main logic elements of the processor. This bus structure, in turn, connects to the VAX UNIBUS arrangement, which provides the communications interface between processor and external connections.

Main Memory Subsystem: The main memory consists of 2 Mbytes of memory array modules, a memory array bus, and a memory controller module. The memory can be expanded to 8 Mbytes to accommodate future growth.

Auxiliary Storage and Peripheral Subsystems: Currently, a minimum of 165 Mbytes of fill on-line storage is anticipated for the minicomputer system. To ensure against any unanticipated or other rapid growth factors, the initial system should be configured with 300 Mbytes or more of auxiliary on-line storage capacity. The storage media used would include:

- Disk Units – To support immediate and future growth requirements, it is recommended that the VAX 750 (or equivalent) be equipped with a UDA50-A Disk-drive Controller supporting an RA81 Winchester Disk Drive Unit (or equivalent). This configuration provides 456 Mbytes of storage capacity.
- Tape Units – The VAX system includes one TU80 data streaming tape unit for making copies of large quantities of data to support off-line storage and manual transfer of files. The TU80 tape unit is a single-density, dual-speed tape subsystem that operates with the UNIBUS and conforms to the ANSI standard for phase encoding (PE) with 1,600 bpi on 1/2 inch, 9 track tape. It has a data transfer rate of 160 kilobytes (Kbytes) at 100 ips.

Input/Output Devices: An LA120 DECwriter is provided as a control/ maintenance terminal.

System Interfaces: Batch files would be transferred between the VAX 750 minicomputer and Sperry 1100/82 mainframe using a communications interface provided by Internet Systems Corporation. This link has been recommended because of its high throughput capability.

The Internet interface equipment will be configured as a front-end processor arrangement between the DEC VAX and Sperry 1100/82 using a Microvax II with Internet V2.6 Sperry Hyperlink proprietary software. The VAX 750 will interface to the Microvax II through an Ethernet connection between the two systems. The Microvax II will provide a block-multiplexer channel connection to the 1100/82. This connection arrangement will provide a communications throughput of 400 Kbps.

Based on the file transfer schedule described in the previous section, the worst case requirement will be to transfer an aggregate of approximately 150 Mbytes of file data. This file data transfer should occur out of normal business hours so as not to cause system response delay for normal RTIS operation. Therefore, batch file transfer of 150 Mbytes must occur within a 7-hour period (midnight to 7 a.m.). At a data rate of 400 Kbps, the time required for this batch file transfer will be 375 seconds.

Communications between terminal users and the VAX will be performed through the DCP/40. The design of this interface is described below.

Telecommunications Considerations: Interfaces between the terminals and the DCP/40 will be the same as those that currently exist. It will be necessary for terminal users to alter sign-on procedures to select the system with which they want to communicate.

Distributed Communications Processor: The interface between the DCP/40 and the VAX 750 will be made using a serial synchronous line module. The DCP/40 will provide a VAX 750 host-to-terminal connection capability through its line modules. The interface between the VAX 750 and DROLS system terminals will be a DEC DUP11 synchronous EIA RS-232/CCITT V.28 connection from the VAX to a line module interface on the DCP/40, through the DCP/40 to the appropriate DCP/40 interface for the accessing terminal. This interface arrangement will support a data transfer rate of 9,600 Kbps.

The 9,600 Kbps interface will also be used for communications between the VAX 750 and the 1100/82 when the input system requires data base inquiries of the DROLS-supported data bases (such as the TR file). The VAX will format a standard DROLS inquiry and communicate through the DCP/40 to the 1100/82 as though it were another terminal. This technique eliminates the need for modifications to the

DROLS software to accommodate minicomputer generated inquiries, since these inquiries can be processed as though they originated from another terminal.

5.2 Support Software Environment.

The hardware/software system installed to provide the required input system processing, must be capable of performing all functions described in the preceding sections of this FD. In general terms, the system must be capable of accepting user inputs, performing the necessary security checks, editing the inputs, updating the data base, and responding to the user. It must also be capable of interchanging data with DROLS, preparing performance reports, and accepting updates to the files used to support the input system operations (such as the MAI files).

The software environment within which these functions are performed includes a combination of:

- Manufacturer-provided support software (operating system, utilities, editors, etc.)
- Specially developed software to perform the necessary security processing, communications interfacing, and formatting functions
- Standard software packages for data base management and spelling checks.

The functions assigned to each of these categories of software depends on the capabilities of the standard products available for the equipment configuration selected for the input system.

5.2.1 Existing System Software Environment.

The existing input system, RTIS, and associated batch programs were developed to operate on the Sperry 1100/82, using a combination of features available from the Sperry 1100 operating system and the associated Transaction Input Processor (TIP). The functions identified above as performed by software packages and specially developed software were designed and programmed by DTIC. With the exception of the data management system (DMS)/1100 software used on the

Sperry 1100/61, general-purpose applications programs such as DBMS software (BASIS) were not used to support this development.

Although the RTIS software was initially programmed to operate as a stand-alone program, it was subsequently modified to operate as an application package under DROLS. Thus, the combined processing capabilities of the Sperry 1100 operating system and DROLS are used to support the RTIS operation.

5.2.2 Defense Research, Development, Test, and Evaluation On-line System Software.

The DROLS processing is described in this section because of the importance of maintaining equivalent processing capabilities for communications interfacing and security processing (user verification) in the replacement input system.

DROLS is a multiactivity reentrant communications and applications software package that provides the executive functions for the control and direction of terminal inputs to the correct applications program. It consists of two primary processing modules, the communications module and the applications module. Data and control are passed between these two modules using a high-speed buffer in which the terminal identifiers, terminal characteristics, function word (defining the function to be performed), the address of the terminal message on the drum file, and various status indicators are stored.

DROLS operation is based on two tables, the scan table and the drum file, which is designated RDRUMSTRINGS. The scan table provides the interface between the two modules, while the drum file serves as the storage medium for the communications messages.

The following sequence of processing steps is performed by DROLS working with the Sperry operating system software:

- Terminal inputs are received from the DCP/40 by the Sperry 1100 and the Message Control Bank (MCB) communications software. That software checks the user's password and security data and stores the input message in a buffer known as the D-Bank. Each terminal input record of the Data Bank (D-Bank) includes 480 words of storage for the input message buffer, 24 words for the MCB packet and auxiliary area, and a buffer area of 320 words for the Sperry FIELDATA text. The FIELDATA storage area is configured for a screen size of 80 characters by 24 lines. That area is adequate for the U100 terminals that have a screen size of 64 characters by 15 lines and the U200 terminals that have a screen size of 80 characters by 23 lines. Addressing limitations have required a modification of the D-Bank from the original technique of fixed buffer areas for each terminal transaction to an indexed addressing scheme that provides additional storage flexibility.

- Following the 1100/MCB processing, the DROLS input routines SGNONA and SGNONS are used to process asynchronous and synchronous terminals, respectively. These routines perform the transaction and program initialization, analysis of input parameters, searching of security tables, and processing of interfacing tables. SGNONA and SGNONS format the TCRTAB record for the the input message and compare the security data in this table with the permanent file of site-control information (ACTAB) which includes data stored for each terminal that is used to verify that all terminal identification entered by the user (and automatically from the terminal), correspond with the data in the ACTAB. The data that is checked include:

  - Routing Identifier (RID)
  - Station Identifier (SID)
  - Device Identifier (DID)
  - Site Sequence Number
  - Terminal connection: in-house, remote, or dial-up
  - User type: DoD or contractor
  - Activity
  - Classification: classified or unclassified
  - Position Identifier (PID)
  - Terminal type.

  The user data are not entered into the files TCRTAB and RDRUMSTRINGS are not created until the user parameters have been verified. When verification has been completed, these files are updated including an entry in TCRTAB that identifies the location of the user's data on RDRUMSTRINGS. The SGNON routines are then exited.

- When the DROLS input processing has been completed, the DROLS applications controller, which is continuously scanning the contents of the file TCRTAB, is responsible for passing control to the appropriate applications program. The applications control section begins with another verification of the user data by comparing the TCRTAB with the ACTAB. This section also accesses user file information that defines the commands the user is authorized to execute. The lists of legal commands include types of searches and sorts, list, display, order, recall, exhibit, transfer, erase, and input commands. The files that can be accessed, including the TR, WUIS, and IR&D files, are also identified for both the main file and the current transaction files of these data bases. If verification of the master control tables is satisfactory, the file is then passed to the appropriate applications programs including the RTIS.

- Five independent application activities can be initiated to process the terminal inputs from the communications modules. Each of the applications activities is programmed using a reentrant code that permits processing of more than one user request at a time. Currently three of the activities (Numbers 1, 2, and 4) are used for processing classified terminals and two activities (Numbers 3 and 5) are used for processing unclassified terminals.

- When the applications processing is completed, the applications controller passes program control to the communications module by setting the appropriate function code in TCRTAB. For example, the communications process can be requested to terminate the session, and in that case, the SGNOFA or SGNOFS routines can be activated to terminate the terminal activity and initialize all files related to the transaction.

In addition to these terminal-processing functions, DROLS maintains the files for site control information and authorized commands/availability, maintains audit trails of all terminal activities, maintains performance records of system processing activities, prints exception reports, and summarizes system operating statistics.

5.2.3 System 11 Software Environment.

The software environment for the System 11 alternative will use the Sperry 1100 operating system and the DMS-1100 and MCB software currently used with the 1100/82. That software will continue to provide the interface communications functions with the DCP/40 TELCON operating system. This "front-end" software design should follow the concept used for the DROLS processing by assembling terminal inputs using message blocks corresponding to the D-Bank concept used by DROLS.

A new executive software package must be programmed to replicate the security processing functions provided by SGNONA and SGNONS. These packages must check the same user and terminal characteristics that are checked by DROLS. The executive software package must be capable of passing all acceptable requests for systems processing to a series of applications packages capable of providing the required input system functions. The use of two data bases similar to the TCRTAB and RDRUMSTRINGS data bases as buffers between the communications software and the applications software should be continued.

The existing multiuser processing technique using re-entrant code, should be replaced with a multiuser concept that takes advantage of the capabilities of the operating system. This revised approach would permit the software to process each terminal input as though it were the only input being received. The operating system would be responsible for scheduling, setup, and execution of all programs. It would also control data base access to prevent contention among users desiring access to the same files.

A number of special applications programs must be developed to support the functions to be provided by the new input system. The required functions include:

- Formatting requests for access to the TR file, which will not be maintained on the System 11 but will continue to reside on the 1100/82. These requests will be formatted as a standard request to the DROLS system for a TR file search.
- Producing form letters for denials and delayed releases and other letters that might be required to document sources.
- Maintaining and updating a problem file of documents that are not acceptable for entry into the current files.
- Providing two alternative screen formats for data entry.
- Printing formatted proof listings.
- Maintaining all authority files and performing the necessary security checks using these files.
- Accepting and performing a spelling check of all document abstracts.
- Running the MAI programs and providing feedback related to the results of the indexing process.

- Converting the current transactions files into the formats required to update the corresponding on-line files maintained by the input system and DROLS.
- Updating all DROLS files automatically or at operator-requested intervals.
- Creating microfiche header tape.
- Performing automated edit-checks and producing list of problems identified.
- Generate AD number labels.
- Preparing random selection of citations for quality-control review.
- Supporting the quality-control process by generating proof listings, extracting documents, providing update/editing screens, and accumulating statistics related to the number of corrections required for each document.
- Generating management reports of pipeline status and production rates.

In addition to these specialized support functions, general-purpose DBMS software will be required to provide the capability to update, search and retrieve, delete, and maintain all data bases supported by the input system. These data bases should include the current transactions versions of the AQ data base, TR, WUIS, and IR&D files. They should also include the same file management capabilities for the files required to support the input process including the terminal and user characteristics, the source header, source hierarchy, fields and groups, DRIT, and MAI files. Security procedures must be implemented as a preprocessing step to the DBMS software to provide restricted access to these files. The DBMS software must be capable of processing abstract data. It must also provide the capability of interfacing with spelling check software, if required, during the input and quality-control processes.

5.2.4 DEC VAX Software Environment.

The DEC VAX software environment is equivalent to that of the System 11. The primary difference between these two systems is the manner in which the communications interface software is implemented. While the Sperry System 11 can be implemented using communications interface modules identical with the 1100/MCB modules currently used on the 1100/82, it will be necessary to develop additional special purpose modules for communicating between the VAX and the DCP/40 for terminal transactions and for communicating between the VAX and the Sperry 1100/82 for file transfers.

The software modules used for communicating with the DCP/40 should provide the capability for exchanging synchronous data blocks with the DCP/40 as though the VAX were a synchronous terminal. Standard communications software is available from DEC to perform this function. The software must be enhanced to perform the functions required to establish the equivalent of D-Bank entries for applications program processing. In addition, this software should be used for directing input system inquiries to the DROLS software resident on the 1100/82, acting as though the VAX were another user terminal.

A second communications interface software system is required to provide the high-speed data transfers between the VAX and the 1100/82. This software would be used in a batch mode to update the files that are resident on both computers.

All other applications software and DBMS software resident on the DEC VAX must be capable of performing the same functions as those described for the System 11.

5.3 Interfaces.

The replacement input system is primarily intended to be a stand-alone process, however some interfaces will be required. These interfaces are listed below:

- On a weekly or biweekly basis TR, WUIS, or IR&D transactions will be transferred from the input system to the DROLS system for updating into DROLS.
- Duplicate check searches executed in the input system must also operate against the DROLS data bases.
- Selected items must be extracted from the DROLS data bases and transferred to the input system for updating.

- Authority files will be maintained both within the input system and within other DTIC applications. For each file, one of the systems will be selected as the master. Updated authority files will have to be transferred after any updates.

- From a hardware standpoint, if the replacement system is placed on separate hardware then appropriate telecommunications and data exchange techniques will be developed to ensure proper communications with the Sperry 1100/82.

5.4 Impacts.

This section describes the impacts of the replacement systems upon the ADP organization. It includes three areas: the organization itself; ADP operations; and effects on the ADP organization during development.

5.4.1 Impacts Upon The Automated Data Processing Organization.

The Telecommunications and ADP Systems Directorate currently has a section within the Systems Design Branch which dedicates a significant portion of its resources to supporting the current input systems. We expect this organization to assume the primary role of supporting the new system. While there will be significant transition activity, the group will not be expected to support both systems as it is intended to discontinue support of the current system.

There will be requirements for this section to learn the data processing techniques and technologies for the new system, which will be significantly different then the current set of assembly language batch programs. During the time which the section is developing its understanding of the new system there will be contractor support in the way of documentation, training, and consulting.

It is expected that the improved documentation and programming features of a DBMS oriented system will enhance the productivity of the group in making future system modifications and developing new data bases.

5.4.2 Impacts On Automated Data Processing Operations.

The replacement system will convert input operations from a primarily overnight batch oriented system to a daytime on-line process. This will reduce the amount of overnight jobs to be run and the number of files to be run. It will increase the daytime computer system load, which will be handled by the recommended acquisition of a minicomputer. In general the number of jobs requiring operator assistance will decrease.

5.4.3 Development Impacts Upon The Automated Data Processing Organization.

The development of the replacement system is intended to be primarily through contractor support. The prime vendor will be responsible for hardware/software selection and acquisition. Additionally, the vendor will be responsible for all specialized programming to customize the standard software to meet DTIC requirements as outlined in this FD. Separate contracting support will be acquired to assist DTIC in facilitating and reviewing the development effort.

DTIC ADP staff will assist the contractors with detailed information regarding integrating the replacement system into DTIC's environment. This will include providing information regarding software and hardware interfaces. It will also provide information regarding training and documentation requirements. DTIC staff may have to provide some programming support to program modules on the Sperry 1100/82 to interface with input system data. Data files will also have to be provided for testing and conversion.

5.5 Failure Contingencies.

Failure contingencies are required to avoid significant losses of data files in the event of processor or peripheral equipment failure. The contingencies discussed here are limited to those failures associated with the minicomputer equipment. It is assumed that contingencies associated with terminal, communications, DCP/40, and 1100/82 failures would remain unchanged.

The primary contingency against data loss must be through the creation of periodic backups of all data bases using the streaming tape equipment included with the minicomputer equipment configurations. Streaming tape has been selected to provide the required backup functions because it is a low-cost, high-capacity, high-speed device designed specifically for mass storage backup.

A tape backup of the supporting files (source header, source hierarchy, fields and groups, etc.) should be created each time the contents of those files are revised. A tape backup of the on-line transaction files should also be created every 24 hours. That tape should remain current until current transactions are transferred to the 1100/82. At that time, the 1100/82 data storage will be considered backup for these files, which can then be reinitialized.

All batch file transfers between computers should be performed by duplicating rather than replacing the existing file that is currently resident on the receiving processor. That procedure will eliminate the possibility of damaging an existing file by simultaneous access to the file while it is being backed up. In addition, the new file should be checked through trial file accesses and verification of file size before it is considered an adequate replacement for the existing file.

As a final step in minimizing the adverse effects of system failures, a complete record of the system operating environment should be maintained on peripheral storage devices so that the system operating configuration can be recreated during start-up following a system failure. In addition, all reports generated by the system including audit trails of terminal activities should be recorded both on disk and tape to protect data from the adverse effects of such failures.

No backup has been provided for terminal transactions that are in progress at the time of system failure. It will be necessary for the user to reinitiate the transaction when the system returns to on-line status.

5.6 Security.

It is important that security measures incorporated in the system processing be equivalent to those that are used with the existing DROLS processing. The requirements that must be satisfied to achieve this level of security include:

- If a minicomputer option is selected, its operating system must have been rated by the U.S. Army Intelligence and Security Command, Counterintelligence Detachment-Defense Nuclear Agency (CID-DNA) at a level of C2 or higher. At the time of this report, Sperry 1100 operating system is the only operating system of the systems considered that have satisfied those requirements. The DEC VAX operating system is currently being evaluated and it is expected to receive a C2 rating in the near future. It is possible that the UNIX Berkeley operating system may also be securable in the future although its status is less certain.

- The security processing activities performed by the DCP/40 should remain unchanged. The DCP/40 currently verifies user RID, SID, and DID of all devices attempting to sign on the system.

- All parameters included in the security checks described in Section 5.2 should be retained. A data base should be maintained that contains an independent set of parameters for each terminal. This data base should be resident on a peripheral storage device connected to the minicomputer. However, it should not be possible to change the contents of this data base from any terminal other than the console terminal of the minicomputer that is located adjacent to the processor.

- Security checking should be performed by the communications software and each applications software routine accessed by the terminal user. The security checks should be performed by comparing the users security parameters with those that are resident in the user data base. In the event that an unauthorized access is requested, the terminal user and DTIC operations personnel should be notified immediately. DTIC personnel should be notified by hardcopy output and console displays.

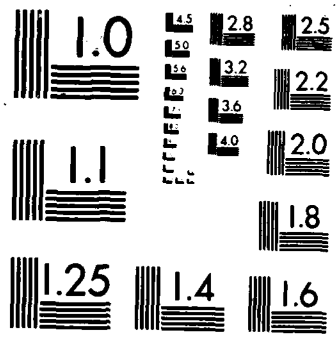
- Audit trails should be maintained for all terminal activities. These trails should identify the terminal, terminal user, time of the activity, duration of the activity, the type of activity, and the applications program used. They should also indicate whether the activity was completed normally or abnormally.

- Summaries of terminal activities should be produced indicating the frequency of each type of activity by terminal, the average duration of the activity, and the maximum duration of the activity. Activities should be categorized according to applications type.

- Exception reports should be produced for all activities in which terminal accesses with durations exceeding a preset time limit are displayed. Duration reports should also be produced for all terminals in which the number of accesses exceeds a preset limit.

AD-A174 522 FUNCTIONAL DESCRIPTION FOR A REPLACEMENT INPUT SYSTEM (U) LOGISTICS MANAGEMENT INST BETHESDA MD D F EGAN ET AL JUL 86 LMI-DL504 MDA903-85-C-0139 2/2

UNCLASSIFIED F/G 5/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

- Wherever possible, accesses from unclassified terminals should be segregated from those from classified terminals. In the multitasking environment described in the software section, this requirement can be satisfied through the storage of two separate sets of communications and applications software on the disk. The unclassified software should be designed in a manner that will not permit it to access classified files. This software would be used by all unclassified terminals. The classified software would automatically be assigned for processing of classified terminals.

- The system must be designed in a manner that prevents its use for software development while the input system is servicing terminal inputs. In a multitasking environment, this will require the physical security of creating a separate library disk that does not contain any development software. In this way, users will be prevented from accessing applications software that is not associated with the input system.

The security procedures should be designed to correspond as closely as possible with the existing DROLS procedures. The ACTAB and the corresponding SCTAB should include identical entries with those used by DROLS. In addition, the report contents and formats should be the same as those produced by the DROLS software. These procedures will minimize the possibility of inadvertently omitting an important security requirement and will simplify the process of reviewing the adherence to existing security requirements.

SECTION 6. COST FACTORS

As indicated in Section 5, a primary objective of the input system improvement program is the replacement of the existing RTIS system as economically as possible. Three approaches have been presented: (1) modification of the existing DROLS system to permit implementation of independent input system processing software incorporating enhanced capabilities and executing on the existing Sperry 1100/82 processor; (2) installation of an additional Sperry System 11 minicomputer; and (3) installation of a DEC VAX 750 minicomputer.

Cost estimates for these alternatives have been developed using cost data provided by equipment vendors, available data related to software packages, and data related to software development costs as a function of program development complexity and size.

As noted in Section 2, the estimated costs of the system improvement will be offset by anticipated efficiencies that result from these improvements.

6.1 Equipment Cost.

The following cost estimates have been developed for the computer equipment and vendor-provided support software of each alternative.

The equipment and system software cost information has been increased by an industry accepted factor of 10 percent to account for the additional costs associated with system integration, cabling, testing, documentation, and training.

- Alternative 1 – Existing System – no hardware costs anticipated

- Alternative 2 – Sperry System 11 Minicomputer

Equipment	Quantity	
Dyadic CPU	1	
Main Memory Storage	1 Megaword	
8436 Disk Storage Unit	2	
2014 Streaming Tape	1	
Controller Card	1	
Block/Byte Channel	1	
Total System 11 hardware cost		$156,000
DCP/40 Line Module (Byte Multiplexer Channel)		4,000
Integration, installation, training, etc.		18,000
Total Hardware Cost		$178,000

- Alternative 3 – DEC VAX 11/750

Equipment	Quantity	
VAX 11/750 CPU		
Main Memory Storage	2 Mbytes	
UNIBUS Disk Adapter	1	
RA81 Disk Unit	1	
TU80 Tape Unit	1	
UNIBUS/Ethernet Controller	1	
DUP11 Comm Interface	1	
LA120 DECwriter	1	
Total DEC VAX hardware		$109,000
Internet Systems Corp. Components		
DEC Microvax	1	
DEC Interface	1	
Sperry Interface	1	

- Alternative 3 – DEC VAX 11/750 (Continued)

Equipment	Quantity
Total Internet hardware	80,000
Integration, installation, training, etc.	18,000
Total Hardware Cost	$208,000

6.2 Software Cost.

This paragraph presents the costs associated with the acquisition of software. The software cost elements include the cost of equipment vendor's system software (including operating system, communications software, and supporting utilities). It also includes the DBMS software cost and the cost of developing the special-purpose processing software required to perform the functions that cannot be acquired with standard packages.

The cost of the vendor-provided software was obtained from standard product catalogs. The cost of the DBMS software is an average cost derived from typical available DBMS offerings.

The cost of the special applications software was developed jointly with DTIC personnel. It assumes that the amount of code required to develop the replacement input system is equivalent to the code developed for the existing system. This is a conservative estimate in that many of the functions performed by the new system would be provided by the DBMS package. However, many new functions will be added to the system and their addition is likely to compensate for the reduction in code resulting from the use of the DBMS. Thus, since the existing package includes approximately 3,000 lines of code and studies have shown that a programming team can produce an average of 6 lines of code a day (including design, coding, testing, and documentation), about 500 man days, or 2 man years, of programming time will be required to complete the support software development. Using an industry average programming cost of $100,000 per man year (including direct salary; overhead; general and administrative cost; profit; and other direct costs of travel, reproduction, etc.) leads to the conclusion that the software development will cost $200,000.

The DEC VAX alternative shows a higher cost for the support software, reflecting the fact that additional code will be required to support the two communications interfaces. In addition, this option will not permit the adaptation of any of the existing RTIS code for the new input system because of the differences in programming languages and operating systems used on the DEC configuration.

Thus the software costs for each alternative include:

- Alternative 1 – Existing System

DBMS	$100,000
Support Software	200,000
Total Cost	$300,000

- Alternative 2 – Sperry System 11

Operating System/Utility Software	$ 18,000
DBMS	100,000
Support Software	200,000
Total Cost	$318,000

- Alternative 3 – DEC VAX/750

Operating System/Utility Software	$ 10,000
DBMS	100,000
Support Software	250,000
Total Cost	$360,000

6.3 Total System Cost.

The individual costs have been combined to include the total system costs. The total system costs include the design cost for each alternative using an estimate of design requirements based on the need to prepare a detailed design of all display screens and functional processes for each alternative. In addition, the minicomputer alternatives require the development of detailed hardware specifications and operating system/utility software specifications that can be used to support the procurement of these systems.

- Alternative 1 – Existing System

Design Cost	$ 45,000
Equipment Cost	N/A
Software Cost	300,000
Total Cost	$345,000

- Alternative 2 – Sperry System 11

Design Cost	$ 75,000
Equipment Cost	178,000
Software Cost	318,000
Total Cost	$571,000

- Alternative 3 – DEC VAX/750

Design Cost	$ 75,000
Equipment Cost	208,000
Software Cost	360,000
Total Cost	$643,000

SECTION 7. SYSTEM DEVELOPMENT PLAN

The DTIC input system study was initiated in April 1985. Analysis of the current procedures was completed in July 1985 and submitted to DTIC in the form of a draft FD in August. In the period of September through December, alternative approaches to the implementation were investigated and evaluated. The results of this evaluation were presented to DTIC in December 1985. DTIC management approved an approach centered upon the use of a commercially available DBMS. It further requested a detailed evaluation of the technical interface between the proposed system and the current DTIC software/hardware environment. The evaluation was to consider options of placing the new software on the current DTIC mainframe or on a stand-alone minicomputer. Between January and March 1986 this evaluation was completed as was the FD of the proposed system. Both the FD and the hardware alternatives are incorporated in this FD.

The next stage of the system development plan is for DTIC management to select a hardware approach and initiate the development of a detailed system specification. The system specification will require approximately 4 months to complete, and the completed specification will serve as the basis for the procurement. The procurement process can be initiated in the fourth quarter of Calendar Year 1986, and contract award can occur in early 1987. The amount of time required for the system development and installation will depend on the chosen hardware alternative and the details of the specification, but portions of the system could be expected within a year of contract award.

Implementation procedures will be detailed in the system specification, but a phased implementation of the data bases and auxiliary files can be expected. System implementation will include parallel testing of the new and current system

APPENDIX A

SHARED BIBLIOGRAPHIC INPUT NETWORK SURVEY RESULTS

A study of Shared Bibliographic Input Network (SBIN) user requirements was conducted to parallel the analysis of internal Defense Technical Information Center (DTIC) procedures and requirements. The SBIN analysis was done through a combination of a mail survey and selected on-site visits. The questionnaire, developed and mailed to SBIN users in May 1985, contained more than 80 questions regarding their current SBIN usage and requirements for a replacement system. The questionnaire was sent to 56 sites listed with DTIC's User Services Section. A summary of the 27 responses received is contained in this appendix.

The site visits were conducted between 26 July and 2 August 1985. The sites visited included seven respondents and four nonrespondents to the survey. Additionally, six Information Analysis Centers (IACs) (four of which are active SBIN users) and the Department of Energy Office of Scientific and Technical Information were also visited. A complete list of the sites visited is given in Appendix B. The needs expressed by the nonrespondents and IACs correspond with those received by the survey and are not included in the summary below except where noted with an asterisk.

The summarized results given here provide numerical quantification of the results wherever possible. However, many of the questions in the survey were open-ended in an effort to get the fullest expression of the user attitudes about the system. These questions are either summarized in sentence or list form. Where several similar responses were received, they are combined and the number of responses noted.

Questions 1-5 and 10-11 referred to the specific identity of the respondent and are therefore not included. Questions 16-25 from Part II and all the questions in Part III have been renumbered to reflect an order of most to least positive responses. For example, 22 of 26 respondents indicated that the ability to duplicate-check documents in the Current (in process) File would be a requirement for any replacement system; it is therefore rated first in this list.

Of the 27 respondents, 9 are not currently performing any SBIN entry and another 3 are only appending their own holding symbols. Of those sites visited that did not respond to the survey, two were not currently active for the lack of trained personnel, one was planning to reactivate by fall, and the fourth felt the effort was currently too time-consuming to perform. Of the five IACs visited, two were very active, one was just beginning, and two were not SBIN users. This survey pertains entirely to SBIN input and does not reflect use of the Defense Research Development, Test, and Evaluation On-Line System (DROLS) retrieval system. The sole exception, Question II-18 which asks about user satisfaction with the retrieval system, functions as a "control" to responses regarding the input system.

	NO. OF RSPNES.	HIGH	LOW	MED.	AVG.
6. Size of library staff? – Professionals	27	20	1	2	4.0
– Clericals	27	27	0	4	6.3
7. Size of library collection (No. of titles)?	24	1,700k	2.1k	150k	201k
8. Number of books in the collection?	27	242k	350	19k	33k
9. Number of technical reports (TRs) generated at your center per year?	20	500	12	100	179
12. Number of documents submitted by your center to DTIC per year? – Via Mail	25	415	0	50	96
– Via SBIN	23	415	0	40	84
12a. Percentage of which with classified citations?	21		8%		

12b. Percentage of which are classified documents?	21	25%
12c. Percentage with 1473 attached?	24	96%
12d. Percentage of SBIN citations with hardcopy follow up?	14	96%

13. For those TRs which you generate, is your catalog record kept:

13a. In a traditional card catalog?	YES: 17
13b. In the DROLS data base?	YES: 13
13c. Your library's own computer?	YES: 3
13d. On another computer you have access to?	YES: 7

14. Please describe any library used computer hardware:

Summary: A wide variety of terminals and printers are in use. However, only four libraries are accessing library files on a mainframe computer, while seven libraries are using microcomputers.

15. Please describe any library-used software on the computer (e.g., LS/2000):

Summary: One library is using LS2000, and the rest are using internally developed systems.

II. SBIN USAGE

1. To which DROLS data bases do you input?

Technical Report (TR):	26
Work Unit Information System (WUIS):	6
Independent Research and Development Data Base (IR&D):	0

2. What type of terminal access do you have to DROLS?

Dial-up:	11
Dedicated Unclassified:	9
Dedicated Classified:	16

3. Number of terminals available to SBIN input? (average): 1.4

4. Specify the type of terminals (brand and model):

   Summary: More than 14 different models of terminals were listed; however, 16 of the 36 were Uniscope 100 or 200's.

5. Specify the line-speed of any dial-up terminals connecting to DROLS:

   Summary: one response of 4,800; all others either 300 or 1,200

6. How many people perform SBIN input? – Professional (average): 1.3
– Clerical (average): 0.4

7. Number of people who have had DTIC SBIN training? (average): 1.5

8. Number of people who need new or refresher training (average): 1.3

9. Time required to key one citation? (average in minutes*): 17
* excludes six responses of between 45 sec. and 6 min. to append holding symbols

10. Number of DROLS input sessions per week? (average): 2.6

11. Number of citations entered per year for documents which were NOT generated at your activity center (Phase IV citations):

Summary: 16 responses of none and one response each of 5, 15, 20, 860, 100, and 4,000 plus.

12. Percentage of these citations which were classified: virtually none

13. Do you dupe-check against the current file? YES: 21 NO: 3

14. Do you dupe-check against the DROLS file? YES: 24 NO: 1

15. What input source do you key from?

Document:	11
DTIC Form 41:	5
DTIC Form 1473:	1
Other Form:	2
No Response:	8

Rate the following: (responses listed in order from most to least positive)	Very Good	Good	Average	Very Poor	Poor	Totals
16. DTIC SBIN input training	8	10	2	1	0	21
17. Local terminal reliability	4	13	6	1	0	24
18. Satisfaction with DROLS retrieval	0	19	5	1	0	25
19. Availability of a classified terminal	8	5	1	2	2	18
20. Availability of CRT to input	4	10	4	4	0	22
21. Input system availability	1	10	8	6	0	25
22. Input system response time	1	6	11	6	0	24
23. Input system user-friendliness	0	7	10	5	3	25
24. Staff availability to input	4	1	6	6	5	22

25. Time required to input	0	5	9	9	2	25

26. Are there any especially significant factors which limit your SBIN participation?

- Inadequate staff to perform input. (seven responses)
- Poor coordination or support from command to obtain documents to input. (six responses)
- No limitations. (four responses)
- Inability to use the same record locally as the DTIC record. (three responses)
- Lack of software and hardware.
- General unreliability of DROLS.
- Dupe-check capability; Input subsystem.
- *Turnover in staff makes it difficult to have SBIN-trained personnel available.
- *Difficulties in using dial-up access.

27. What could DTIC do to help increase your SBIN participation?

- Improve user-friendliness/simplify process. (six responses)
- Improve response time. (two responses)
- Upgrade edit procedures. (four responses)
- Provide local training. (three responses)
- Input form to reduce keying time. (two responses)
- Make DROLS reliable; make system compatible to transfer to other systems.
- User designed output reports; reduce display of edit statistics.
- Eliminate errors due to punctuation and capitalization.
- Instant dupe checking.
- Visit site to coordinate data base with command and library.
- Implement the Local Automation Model (LAM) as soon as possible; LAM capability to handle both MARC and Committee on Scientific and Technical Information (COSATI) formats.

- Ability to tag multiple documents at one time.

- Access during non-regular duty hours.

- Free documents.

- We are going full blast now.

- *Provide incentives for SBIN input such as credit for search time or document ordering.

28. What are the best features of SBIN?

    - Ability to maintain local collection through use of holding symbols. (six responses)

    - Complete records with text available on-line; helpful DTIC personnel.

    - Provides a master file for storing and retrieving studies.

    - Organizational control of subject analysis.

    - Ability to announce reports quickly (two responses); ADD program.

    - Relatively easy to add material.

    - Reasonable cost and good technical support from DTIC.

    - Local control of distribution if desired.

    - Shared cataloging concept.

    - User friendly software allows information to be announced more quickly.

    - None.

29. What are the worst features of SBIN?

    - Difficulty of use, lack of formatted screens, default inputs, correlation to Form 41, etc. (six responses)

    - Having to wait for listings to arrive by mail. (three responses)

    - Losing data due to system problems. (three responses)

    - Inability to perform batch transmissions. (two responses)

    - Response time.

    - Information seems more relevant to Technical Abstract Bulletin (TAB) production than cataloging.

    - No on-site training.

- Inconsistency of what system will accept as accurate.

- Too many variations in open ended terms.

- Time required to coordinate with command to convince them to release documents for SBIN input.

- All.

- *Inability to global edit IAC subject terms.

30. What if any, changes would you like to see in the information kept in DROLS?

    - Allow agencies other than SBIN sites to append holding symbols and display in order to improve local access. DTIC delivery is at times too slow.

    - Display holdings symbol.

    - Capability to append holdings symbol for hardcopy and fiche for the same document without overlaying the other.

    - A field for subject headings and a holdings field for networking.

    - A save procedure and a format field.

    - Addition of retrieval of documents based on source code for sponsor of document

    - Have fields 8a-c of Form 1473 available on DROLS.

    - Major revision of COSATI.

    - Alter DROLS statement regarding non-availability of a document to read: "not available from DTIC yet" (adding the word yet).

    - Eliminate source codes; consistent report number format; group like fields together.

    - Make corporate authors searchable.

    - Adequate.

31. Other comments?

    - On-line editing.

    - Cataloging time too long due to lack of correlation between Form 1473 and input system.

    - Biggest problem is lack of staff time and duplication of effort from local cataloging.

    - Input system must be linked with LAM; this is third questionnaire in less than a year on similar topics.

- Free text searching of abstract; add field for controlled document number for local use; qualifying results of search on general term is too slow.

- Display of holding symbol.

- Retrieve holding symbol by date.

- Ability to associate missing TRs from terminated contracts through DROLS searching.

- Lack of quality control especially in descriptors and identifiers.

III. Rate the following features for a replacement SBIN system (responses are listed in order).

Ability to:	Required	Desirable	Not Needed
1. Search and dupe-check the current file.	22	3	1
2. Search and dupe-check other DROLS files.	19	5	1
3. Produce hardcopy print-outs.	16	9	1
4. Display authority files (DTIC Retrieval and Indexing Terminology [DRIT] etc.) on-line.	15	11	0
5. Generate production statistics.	15	10	1
6. Have citations immediately updated to DROLS.	15	7	1
7. Enter unclassified citations to classified documents.	18	3	5
8. Have edit-checks of one field based on the data in another	15	9	2
9. Modify citations which you already have in DROLS	16	6	4
10. Modify your citations in the current file.	14	7	4
11. Maintain a TR acquisitions file.	11	14	1
12. Perform local entry off-line then batch transmit to DROLS.	10	14	2
13. Enter non-DTIC fields for local use.	9	14	2
14. Access systems other than DTIC through the terminal.	9	13	2
15. Enter classified citations on-line.	12	6	8
16. Use more than one terminal simultaneously.	8	13	4

17.	Define edit-checks for local use.	7	16	3
18.	Maintain a TR pre-publication file.	4	20	2
19.	Modify citations input by someone else.	6	11	9
20.	Access to DROLS beyond current operating hours (0800-1930 ETZ).	4	9	12

APPENDIX B

SHARED BIBLIOGRAPHIC INPUT NETWORK SITE VISITS

Shared Bibliographic Input Network (SBIN) site visits were conducted from 26 July through 2 August 1985. The sites visited included nine SBIN sites (five of which were respondents to the survey) and six Information Analysis Centers (IACs) (four active SBIN participants). The trip was conducted jointly by representatives of Logistics Management Institute (LMI) and Defense Technical Information Center (DTIC). Discussions with the site personnel centered primarily on the input system; however, other topics regarding relations to DTIC were also covered. The complete list of visits is given below. Three organizations visited are not SBIN members. The Air Force Geophysics Laboratory and the Foreign Technology Division of the Air Force Institute of Technology are major DTIC contributors without being SBIN members. The Department of Energy Office of Scientific and Technical Information has a mission very similar to DTIC. The visit to them was for the purpose of comparing their input technology to DTIC's.

I. SBIN Sites

1. Naval Coastal Systems Center
   Panama City, Florida

2. Air Force Engineering and Services
   Tyndall AFB, Florida

3. Air Force Armament Laboratory
   Eglin AFB, Florida

4. Redstone Scientific Information Center
   Hunstville, Alabama

5. Air Force Communications Command
   Scott AFB, Illinois

6. Army Aviation Command
   St. Louis, Missouri

7. Air Force Human Resources Laboratory
   Wright-Patterson AFB, Ohio

8. Air Force Wright Aeronautical Laboratory
   Wright-Patterson AFB, Ohio

9. Army Materials and Mechanics Laboratory
   Watertown, Massachusetts

II. Information Analysis Centers

1. Guidance and Control IAC
   Illinois Institute of Technology
   Chicago, Illinois

2. Manufacturing Technology IAC
   Illinois Institute of Technology
   Chicago, Illinois

3. Metal and Ceramics IAC
   Battelle Memorial Institute
   Columbus, Ohio

4. Tactical Technology IAC
   Battelle Memorial Institute
   Columbus, Ohio

5. Aircraft Survivability IAC
   Anamet Laboratories Inc.
   Wright-Patterson AFB, Ohio

6. Survivability/Vulnerability IAC
   Booz, Allen & Hamilton Inc.
   Wright-Patterson AFB, Ohio

III. Other Visits

1. Department of Energy Technical Information Center
   Oak Ridge, Tennessee

2. Air Force Institute of Technology
   Foreign Technology Division
   Wright-Patterson AFB, Ohio

3. Air Force Geophysics Laboratory
   Hanscom AFB, Massachusetts

APPENDIX C

DEFENSE TECHNICAL INFORMATION CENTER INPUT SYSTEM DATA PROCESSING FLOW CHARTS

On the following pages are copies of Defense Technical Information Center (DTIC) generated flow diagrams of the path of technical report (TR) input data from its input on cathode ray tubes (CRTs) using a Remote Terminal Input Subsystem (RTIS) through its update into the TR data base on a Defense Research Development, Test, and Evaluation On-Line System (DROLS). It also includes auxiliary operations such as updates of the source header and inventory files. These diagrams are not current but are generally reflective of the current operation. As discussed in the text, this flow will be significantly altered by the software normalization effort currently being undertaken by the Telecommunications and Automated Data Processing (ADP) Systems Directorate. They are shown here only to illustrate the complexity and the extent of the data processing operation required to support TR processing. The bulk of the programs shown are currently written in Sperry-UNIVAC assembler language. A major objective of the normalization effort will be to convert them to Common Business Oriented Language (COBOL). Almost all of the files shown are sequential tape files, many of which are multiple reels in length. Last, with the exception of the Remote Terminal Input Subsystem (RTIS), the programs are all batch jobs run over-night. Section 2.3.9 describes the background environment of DTIC's data processing operation.

FIGURE C-1. CURRENT INPUT SYSTEM FLOWCHART

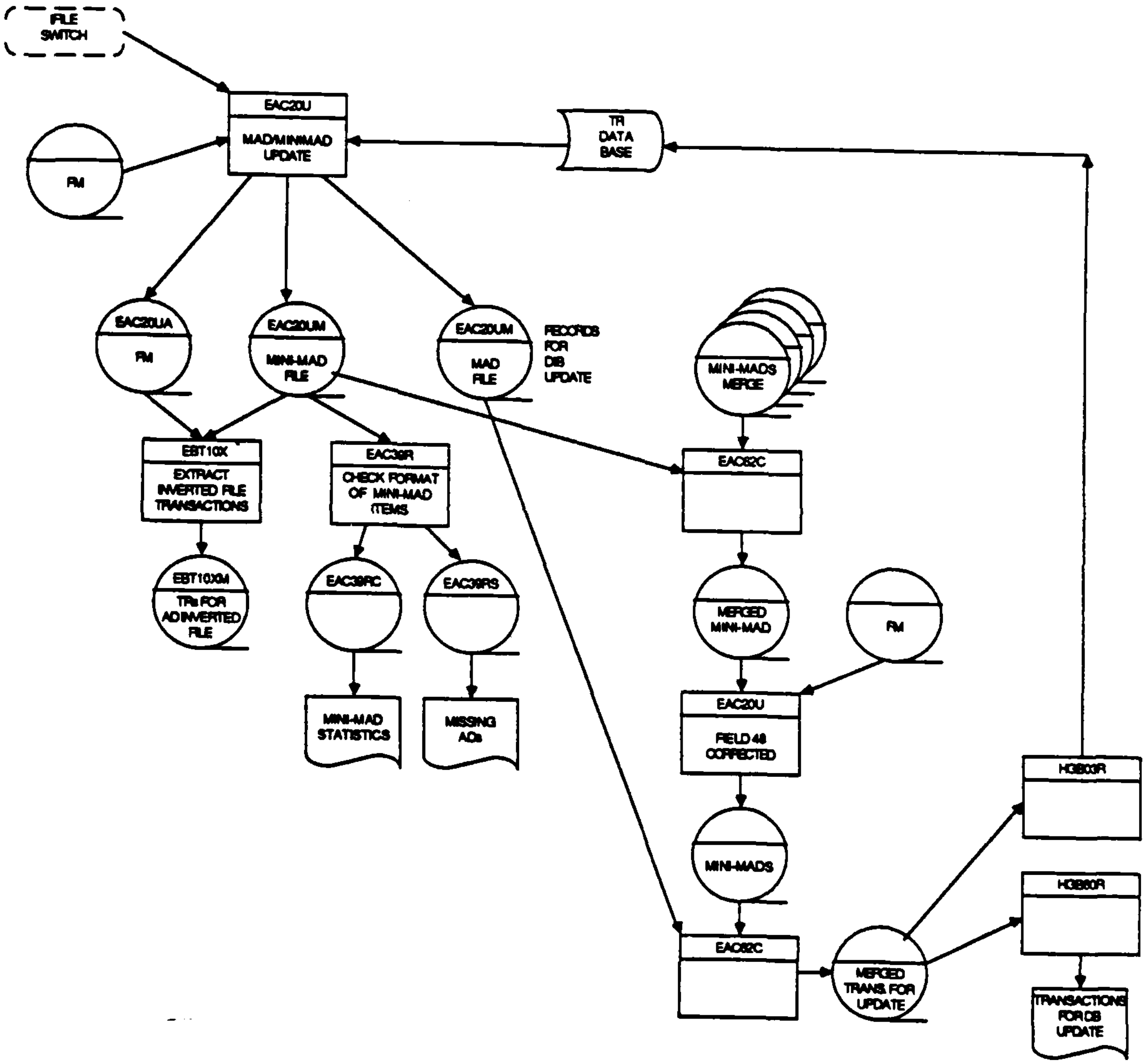


FIGURE C-1. CURRENT INPUT SYSTEM FLOWCHART (CONTINUED)

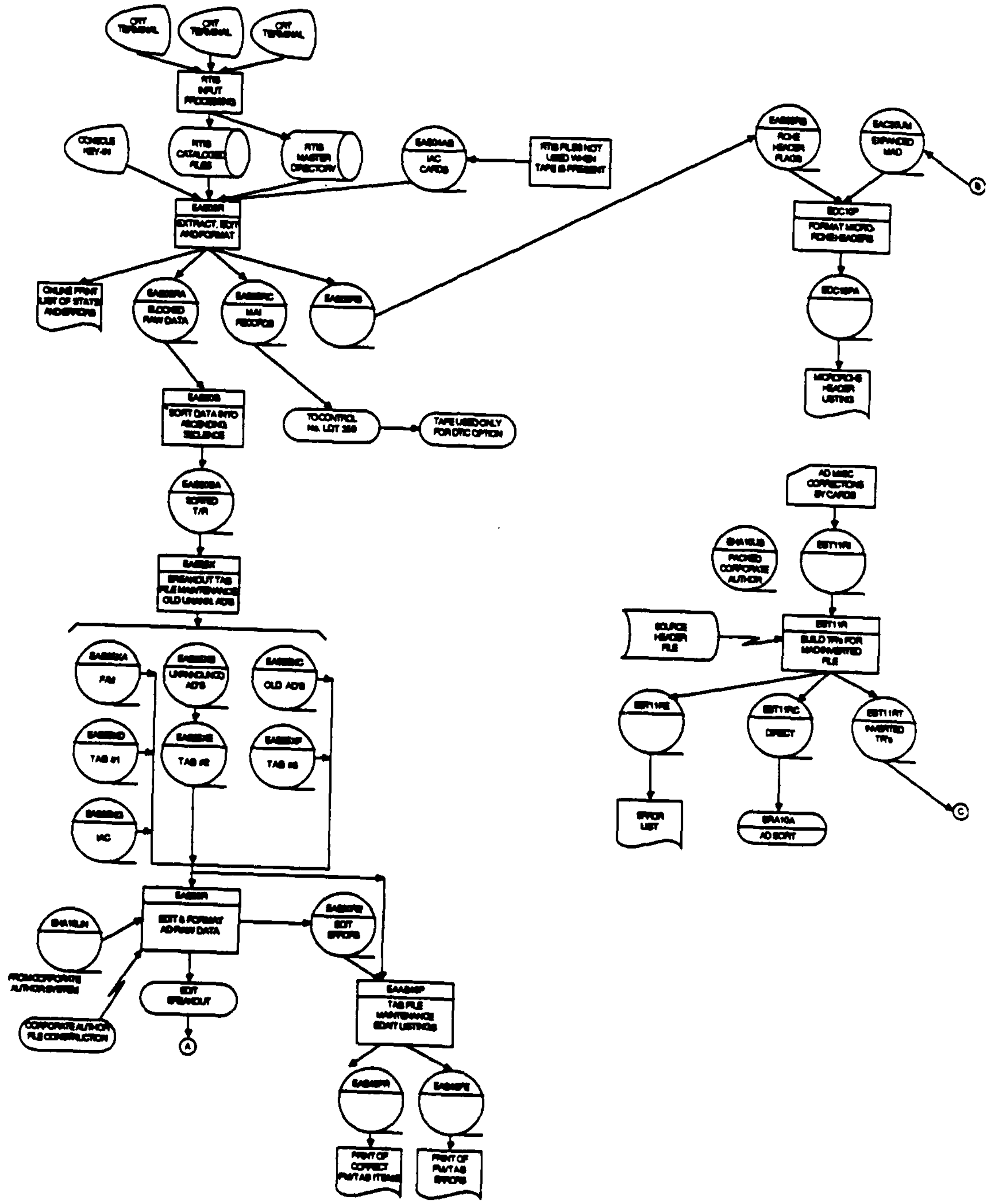


FIGURE C-1. CURRENT INPUT SYSTEM FLOWCHART (CONTINUED)

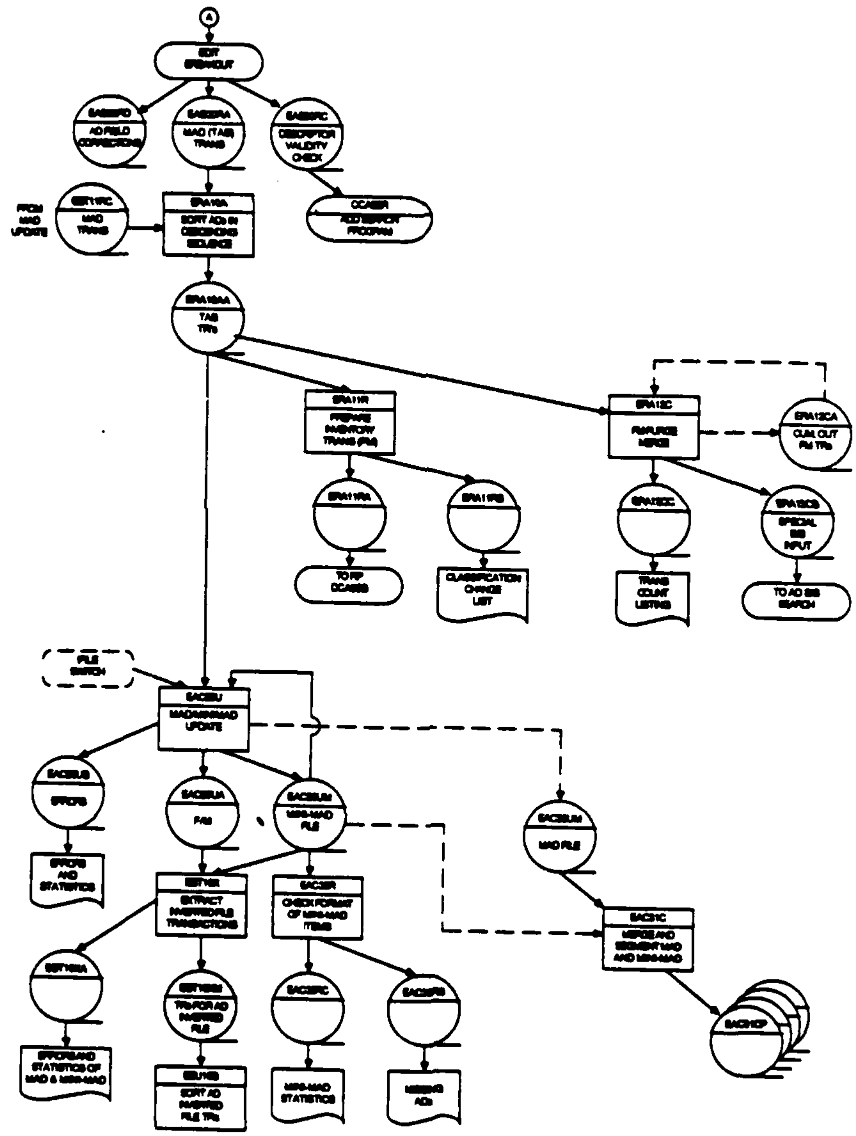
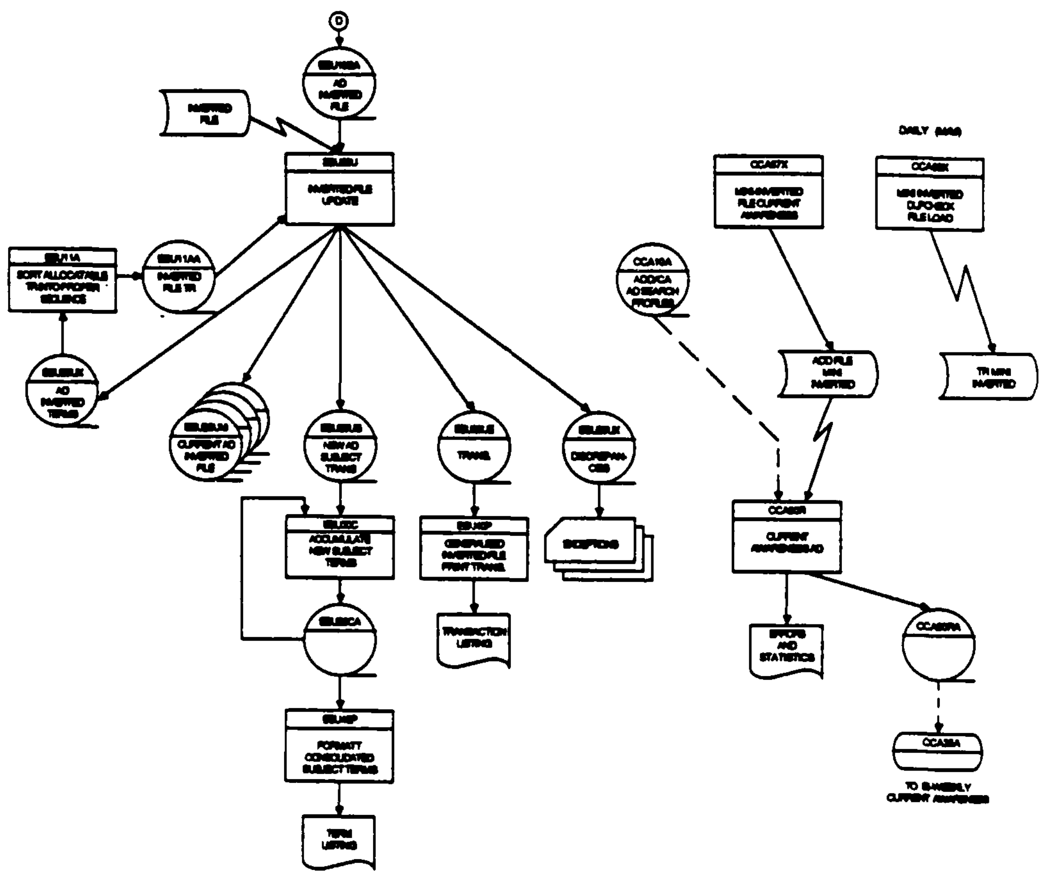


FIGURE C-1. CURRENT INPUT SYSTEM FLOWCHART (CONTINUED)



APPENDIX D

DEFENSE TECHNICAL INFORMATION CENTER INPUT PROCESS FLOW CHARTS

This appendix contains detailed flowcharts which illustrate the flow of document processing described in Section 2.3. The index below correlates each chart to the relevant portion of Section 2.

- Figure D-1: Flowchart Symbol Definition – provides an explanation of all symbols used in the flowcharts.

- Figure D-2: Acquisition Processing – For technical report (TR) documents described in Section 2.3.1. The first page describes acquisition processing from the receipt of information regarding a potential acquisition through contacting the source to obtain the document.

  The second page describes acquisitions' response to each of four possible outcomes to their attempt to obtain the document. Note that the flowchart shows notifying a user of the results of the acquisition effort. This step is deleted if Acquisitions initiated the attempt themselves.

- Figure D-3: TR File Processing – This Chart defines the flow of TR documents through the "pipeline" as described in Section 2.3.2 through 2.3.6.

  The first page describes the flow from receipt of a document in the mail room (Section 2.3.2), through selection (2.3.3), and into descriptive cataloging (2.3.4).

  The flow continues from duplicate checking within descriptive cataloging through checking a document in the acquisition (AQ) data base file for any previous acquisition effort. All of this is described in Section 2.3.4.

  The third page completes the cataloging portion (2.3.4) and illustrates the different indexing processes (2.3.5). Note, that for unclassified documents the data and the document follow a different path for a short time.

  Page four completes the indexing portion (2.3.5) with the keying of indexing terms and the transfer of documents to verification (2.3.6). This chart also shows remote Shared Bibliographic Input Network (SBIN) entry (2.3.7).

  The last chart shows the completion of the verification work (2.3.6). It also illustrates the file maintenance operations of the verification group described in Section 2.3.6. This chart completes the processing of documents for the purposes of this study. Subsequent processing, not described includes updating of the master TR data base and production of the various output products. Details of operations by DTIC-S in support of the input process are shown in the flowcharts in this appendix and described in Section 2.3.9.

- Figure D-4: Source Header and Hierarchy Processing – Displays details of establishing and maintaining the headers and hierarchy. These are shown in brief in Figure D-4 and described in Section 2.3.4. Each is shown on a separate page within this figure.

- Figure D-5: TR Security Change Processing – This chart illustrates processing by the TR Verification Section as described in Section 2.3.6. The first chart initiates security processing from either receipt of a change notification from the source or the routine printing of an Autodog report. The chart shows the three parallel processes which occur. Note that users are only notified for upgrades. The concluding steps in the process are shown on the second page.

- Figure D-6: Work Unit Information System (WUIS) and Independent Research and Development Data Base (IR&D) Processing – The information in these charts illustrates the descriptive text in Section 2.3.8.

  The first chart describes the initial receipt of material in any of three different media (tape, SBIN, and hardcopy) and the initial processing of it. Note that when tapes are converted and a listing is reviewed, if the tape is acceptable processing continues, otherwise the tape is returned to the submitter.

  The remainder of the process is shown on the second page which describes the computer processing of WUIS/IR&D data and the review of the results. For purposes of clarity these charts include the actual updating of the Defense Research Development, Test, and Evaluation On-Line System (DROLS) data base for WUIS/IR&D even though it is not within the proposed scope of the replacement input system.

FIGURE D-1. FLOWCHART SYMBOL DEFINITION

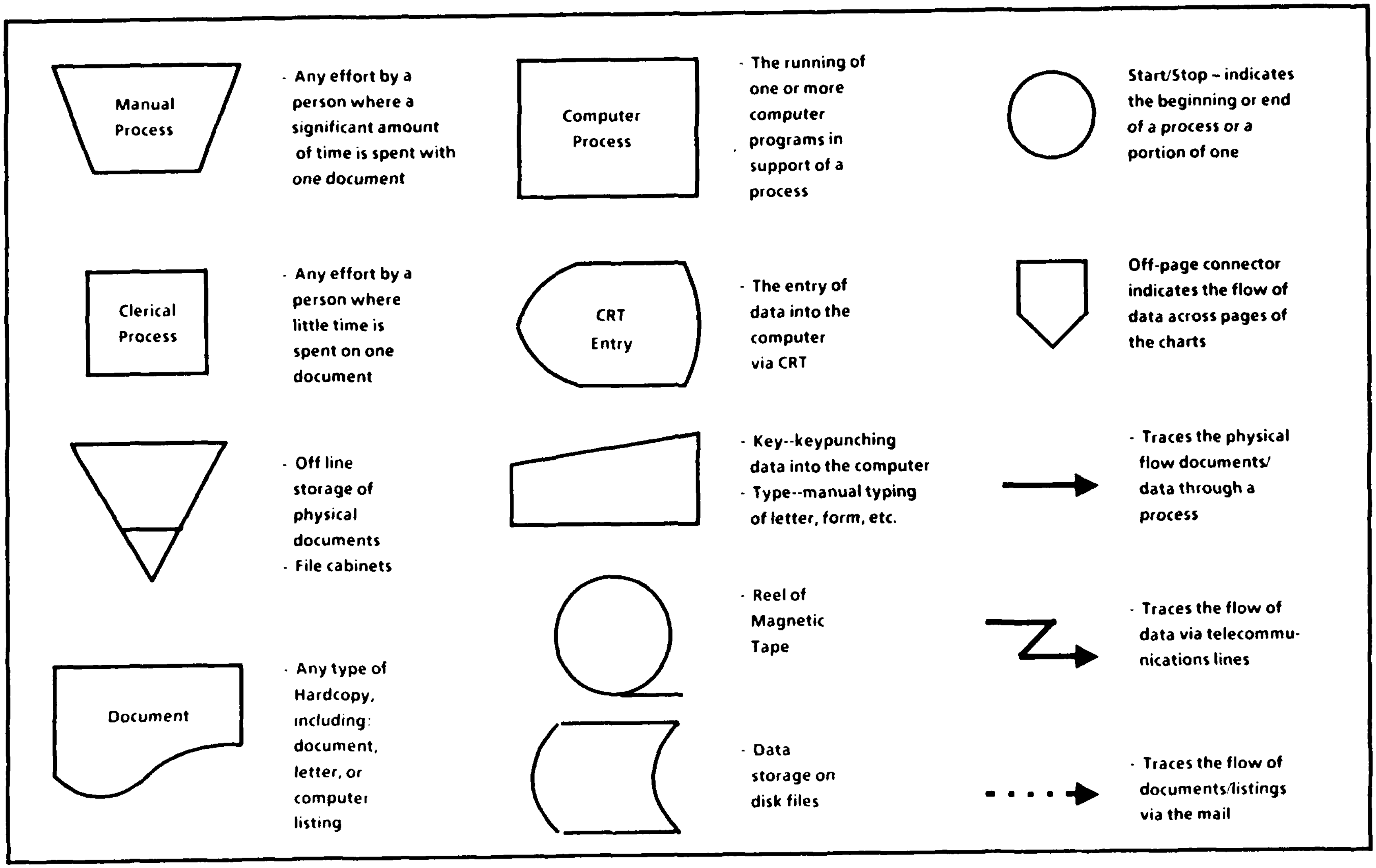


FIGURE D-2. ACQUISITIONS FLOWCHART – CONTACT PROCESSING

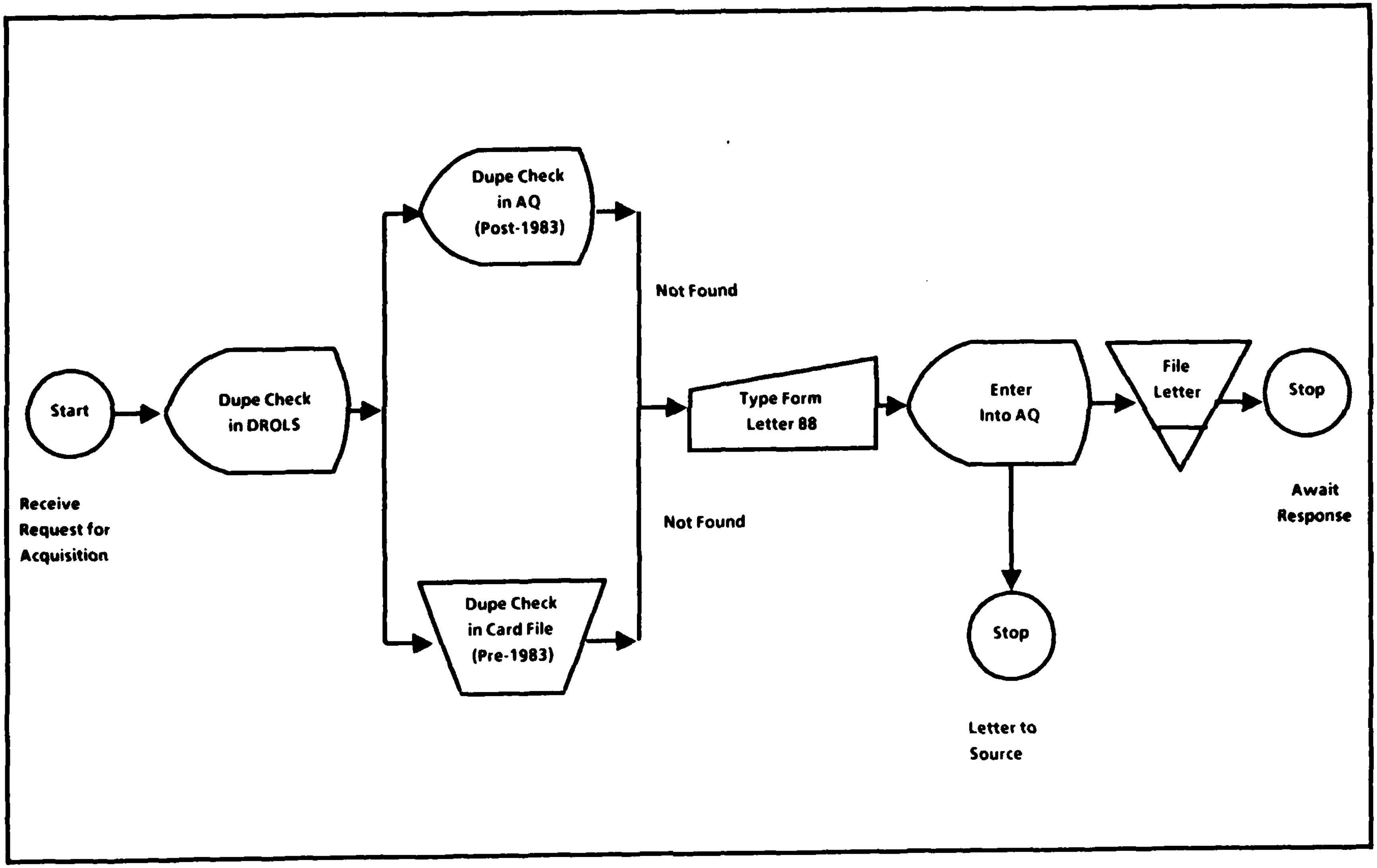


FIGURE D-2. ACQUISITIONS FLOWCHART – RESPONSE PROCESSING (CONTINUED)

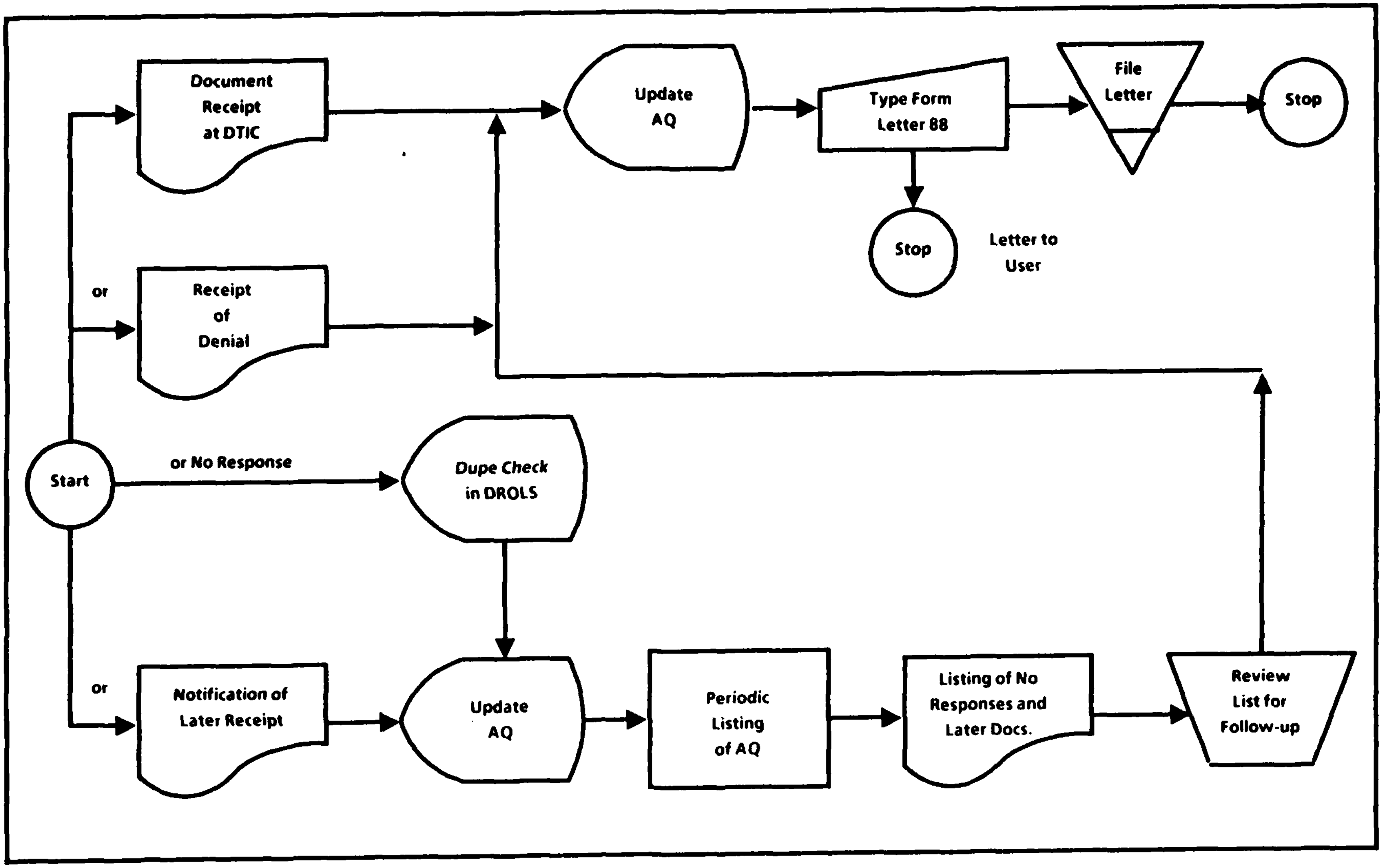


FIGURE D-3. <u>TR FILE FLOWCHART</u>

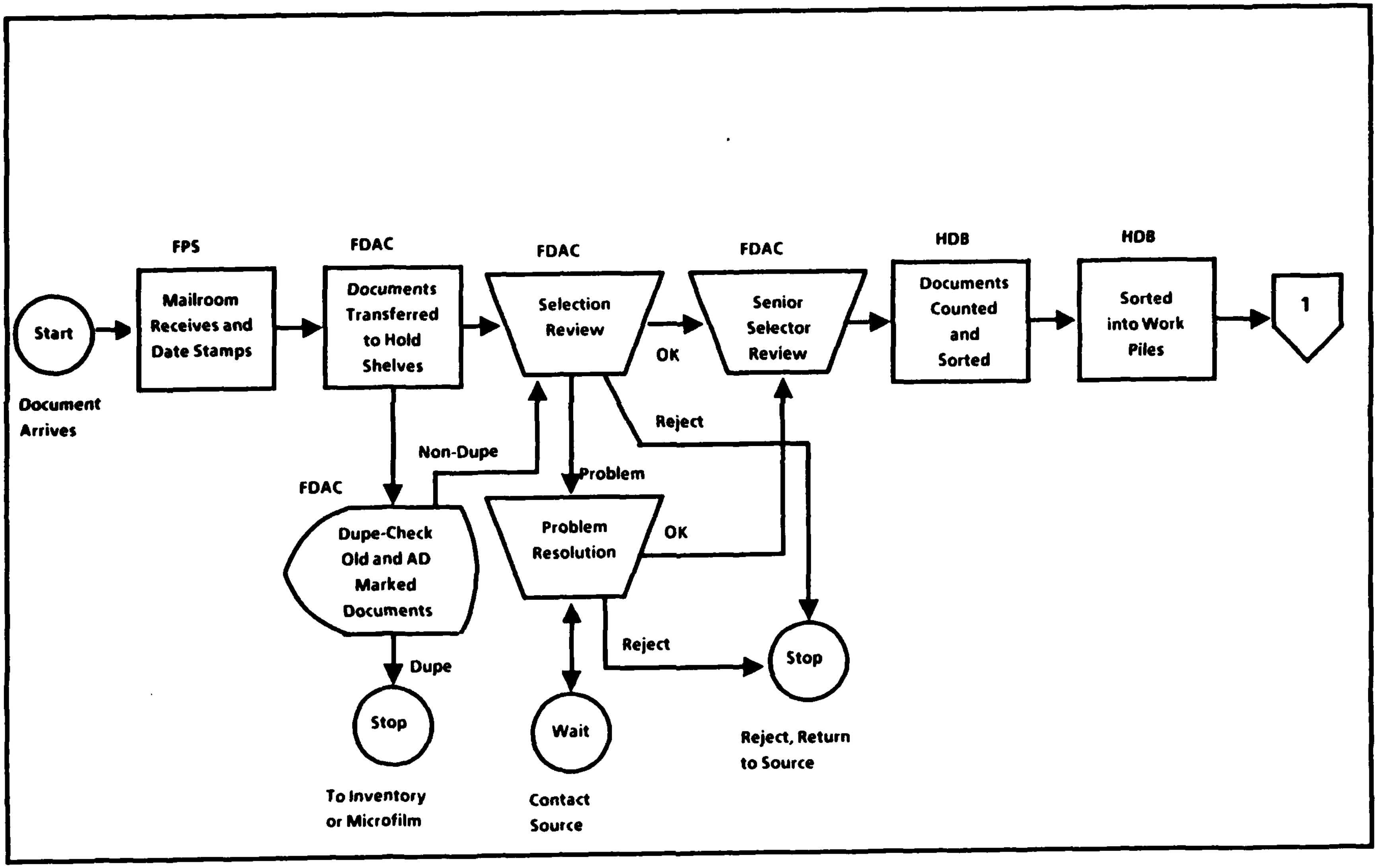


FIGURE D-3. TR FILE FLOWCHART (CONTINUED)

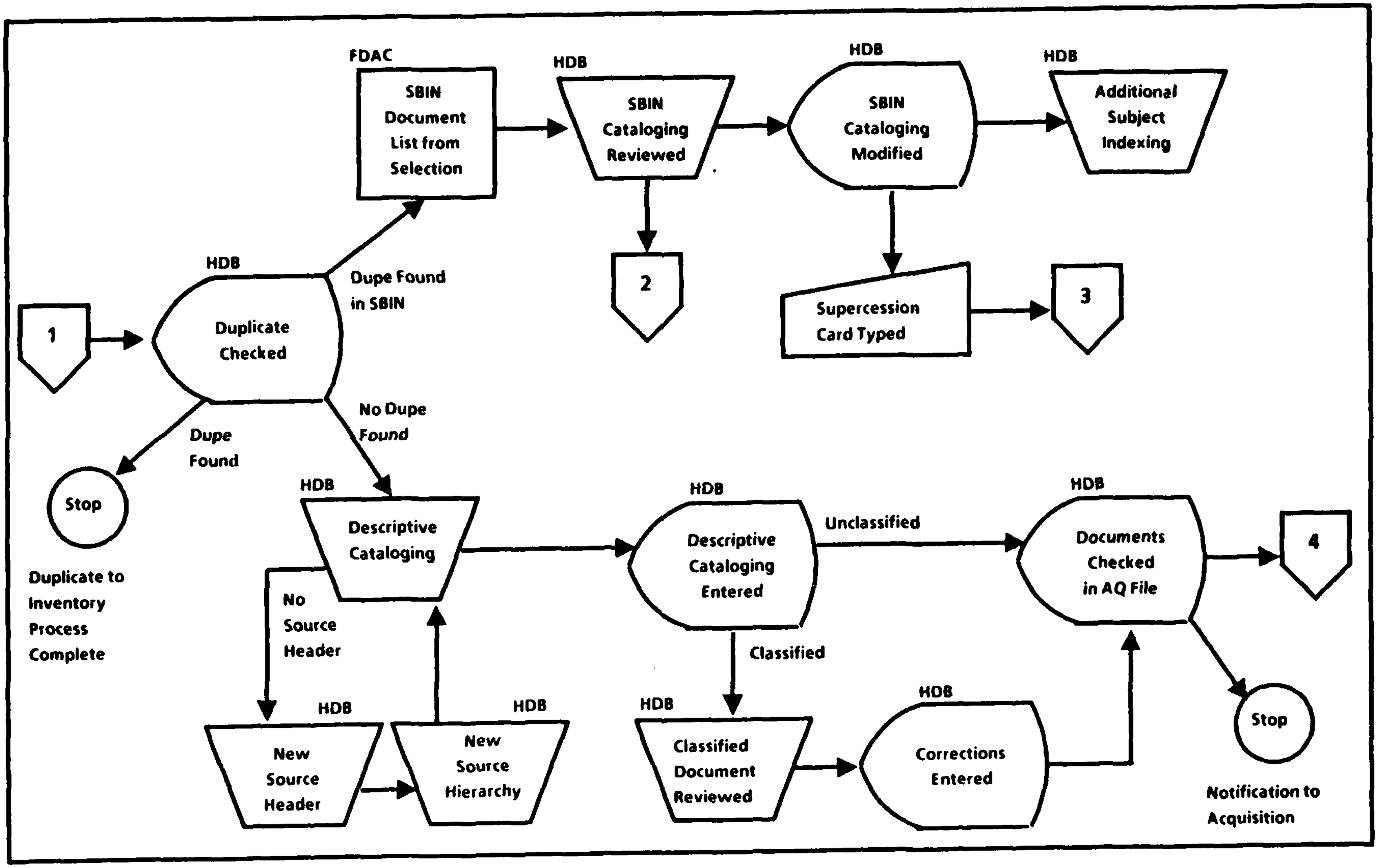


FIGURE D-3. TR FILE FLOWCHART (CONTINUED)

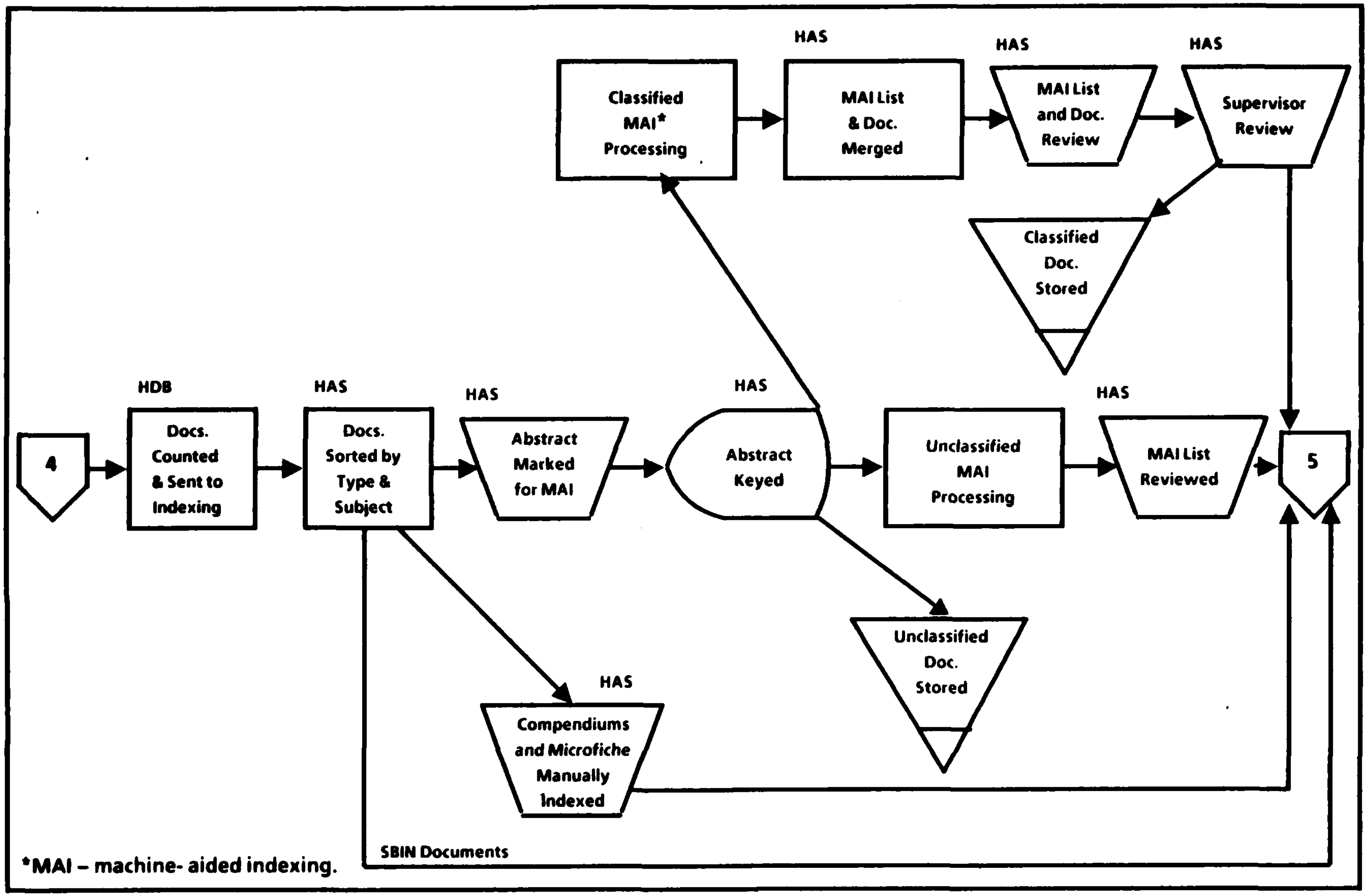
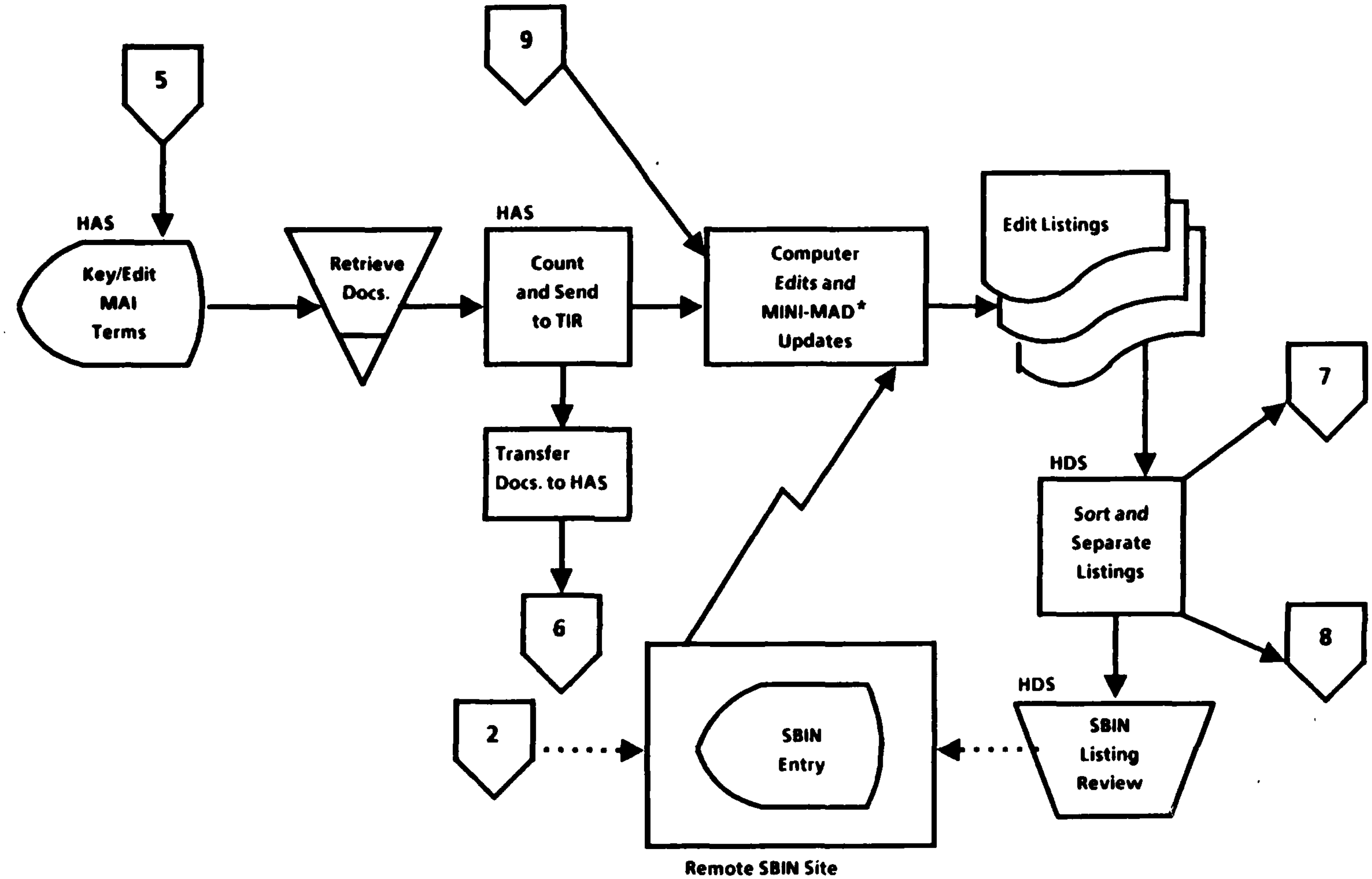


FIGURE D-3. TR FLOWCHART (CONTINUED)



*MINI-MAD – miniature master AD file

FIGURE D-3. TR FLOWCHART (CONTINUED)

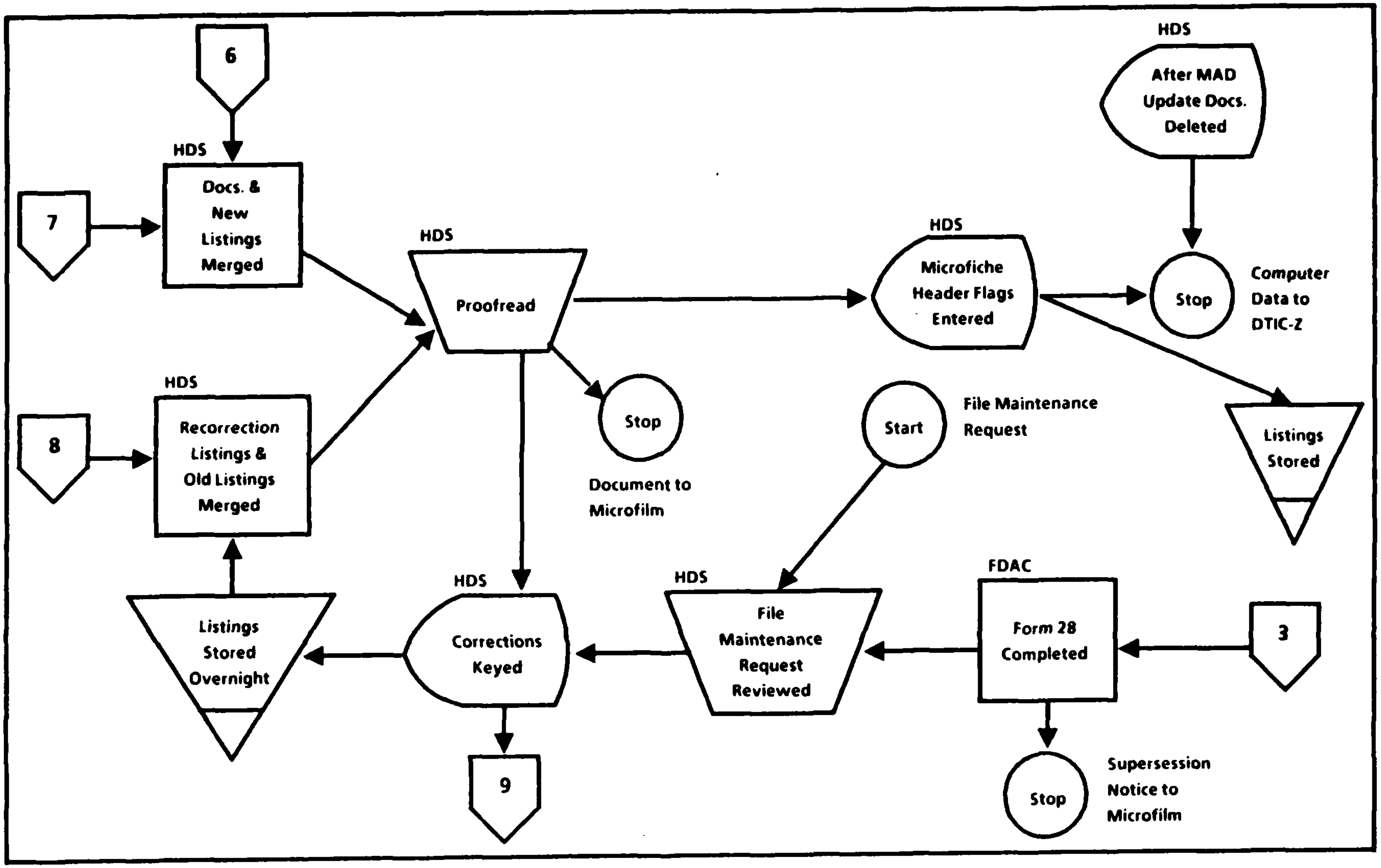


FIGURE D-4. SOURCE HEADER AND HIERARCHY PROCESSING FLOWCHART

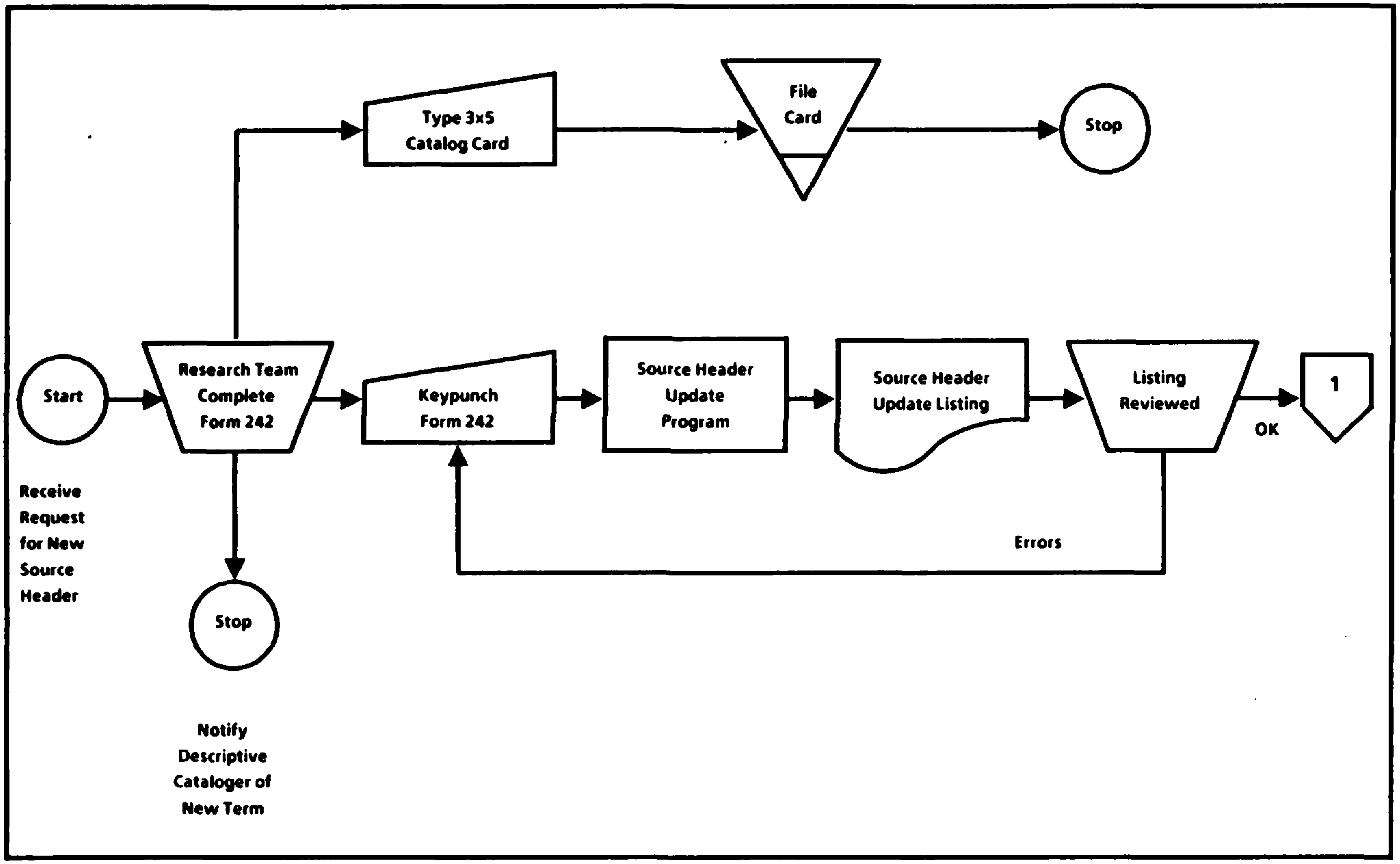


FIGURE D-4. SOURCE HEADER AND HIERARCHY PROCESSING FLOWCHART (CONTINUED)

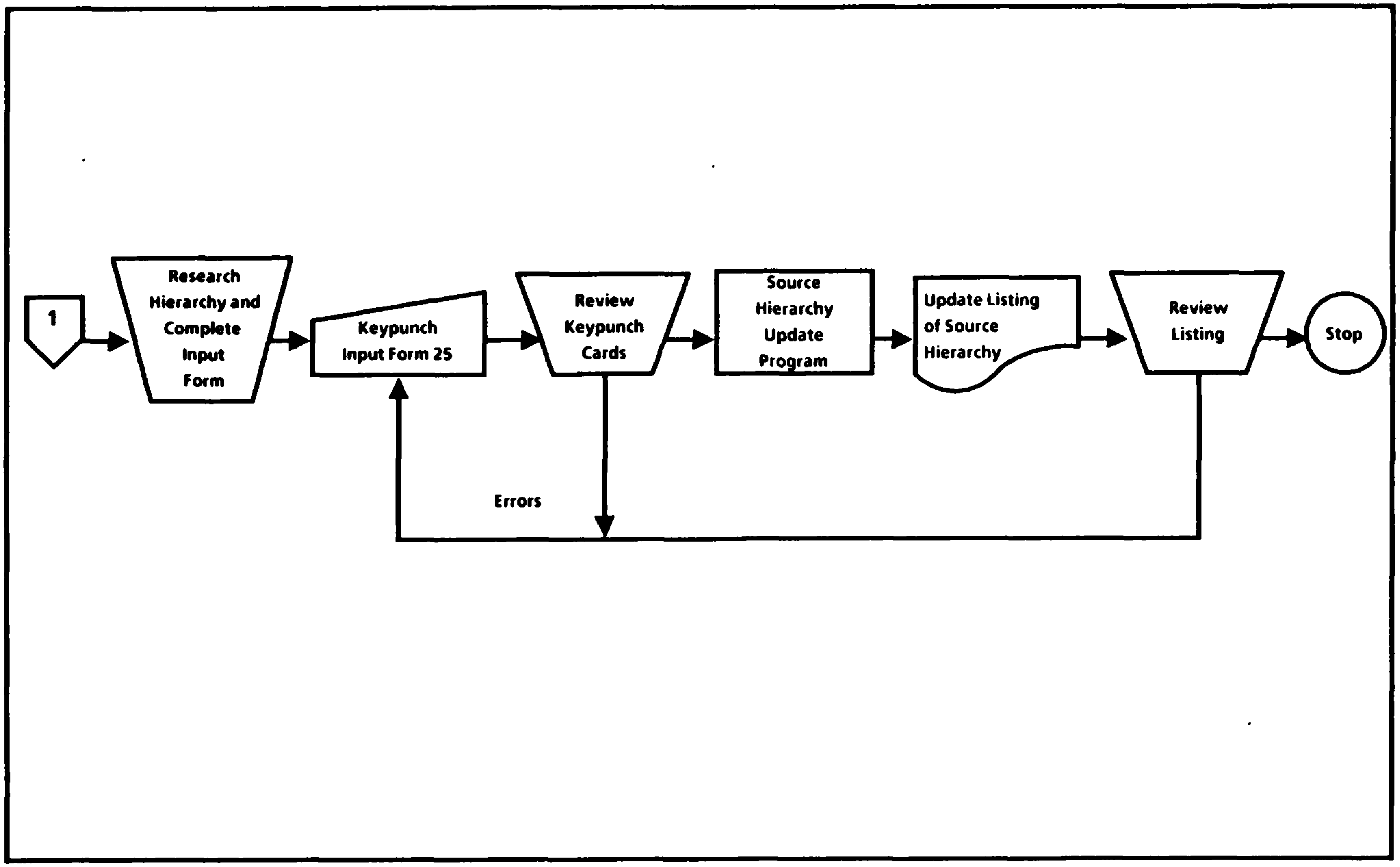


FIGURE D-5. SECURITY CHANGE FLOWCHART

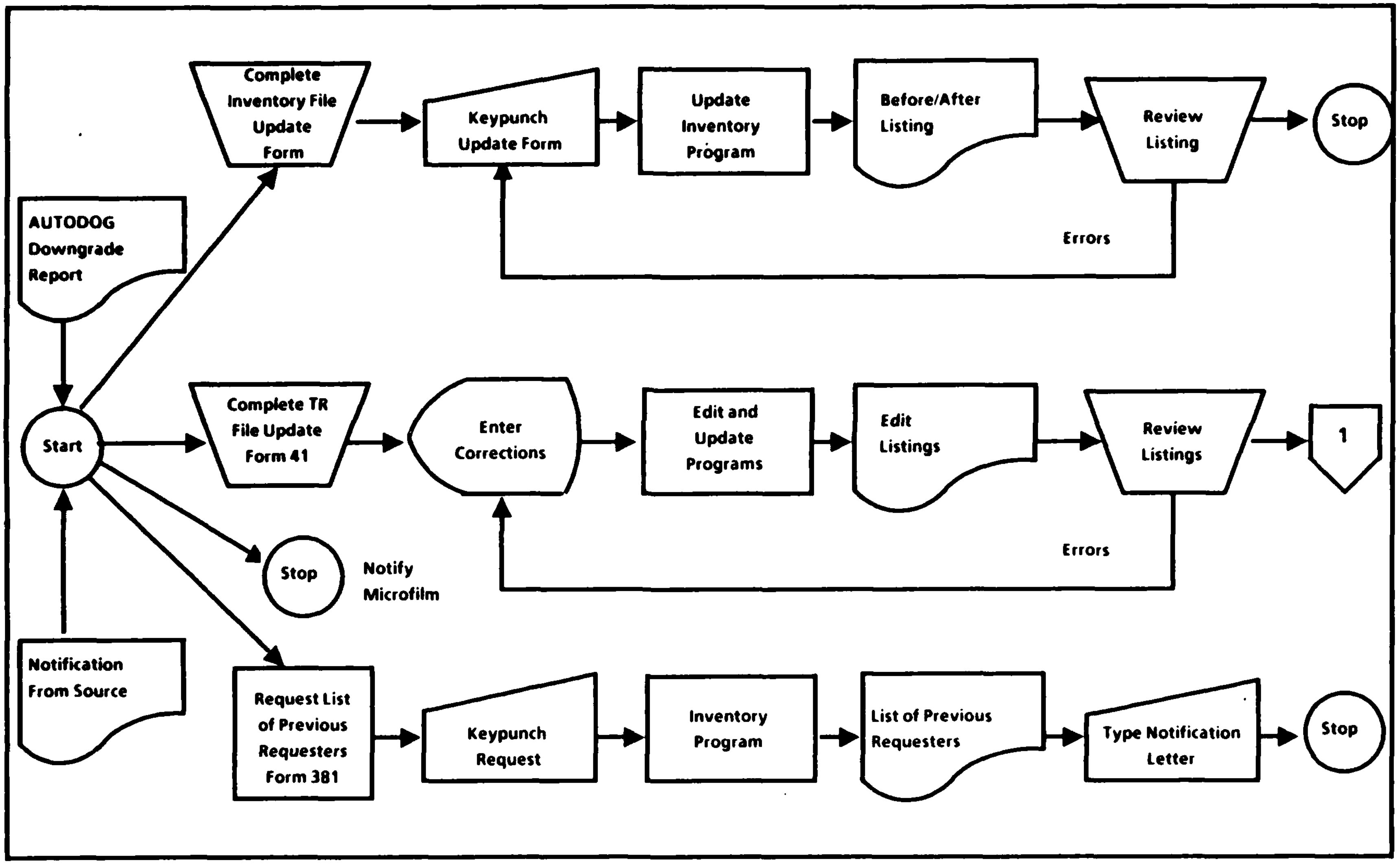


FIGURE D-5. SECURITY CHANGE FLOWCHART (CONTINUED)

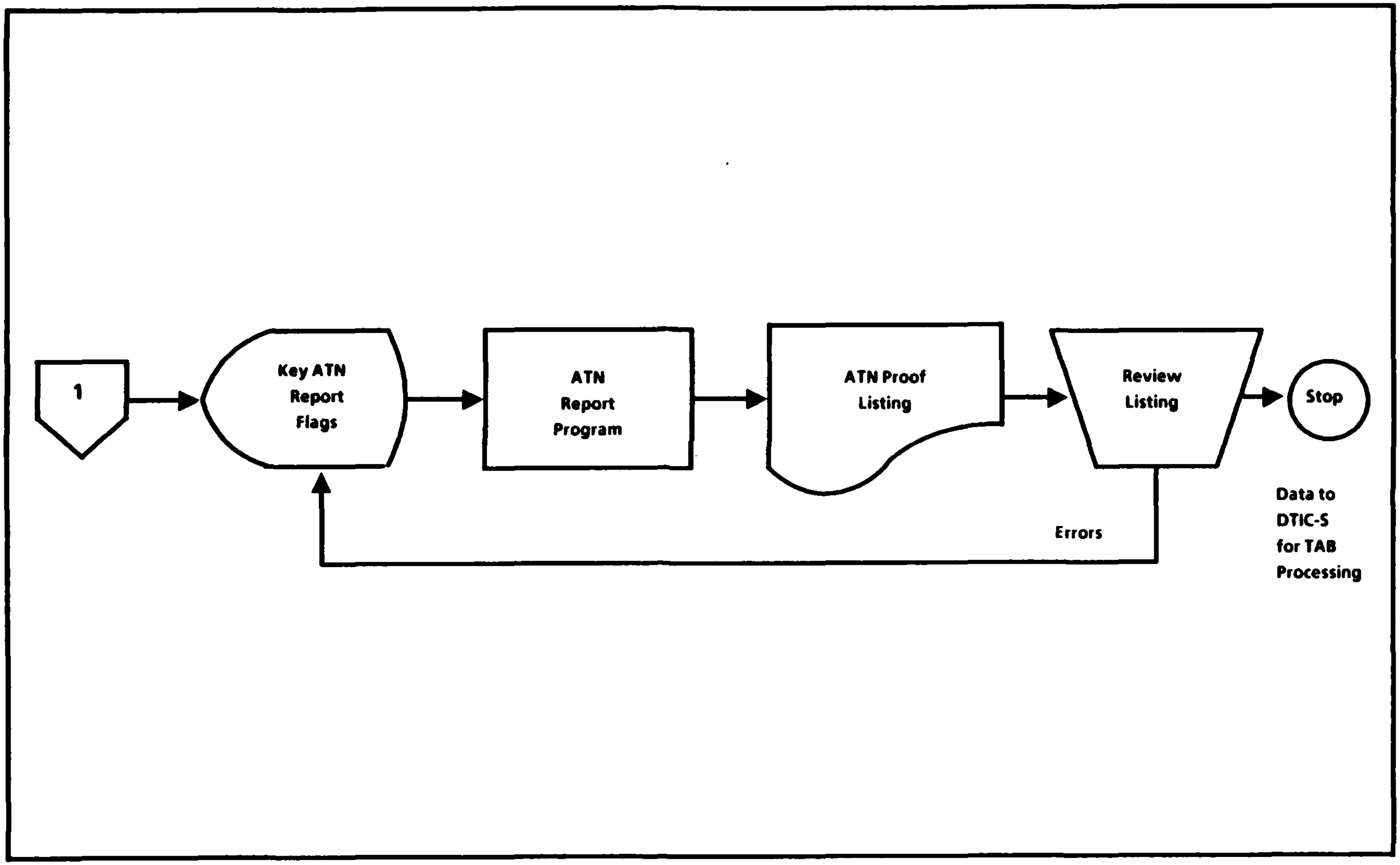


FIGURE D-6. WUIS/IR&D PROCESSING FLOWCHART

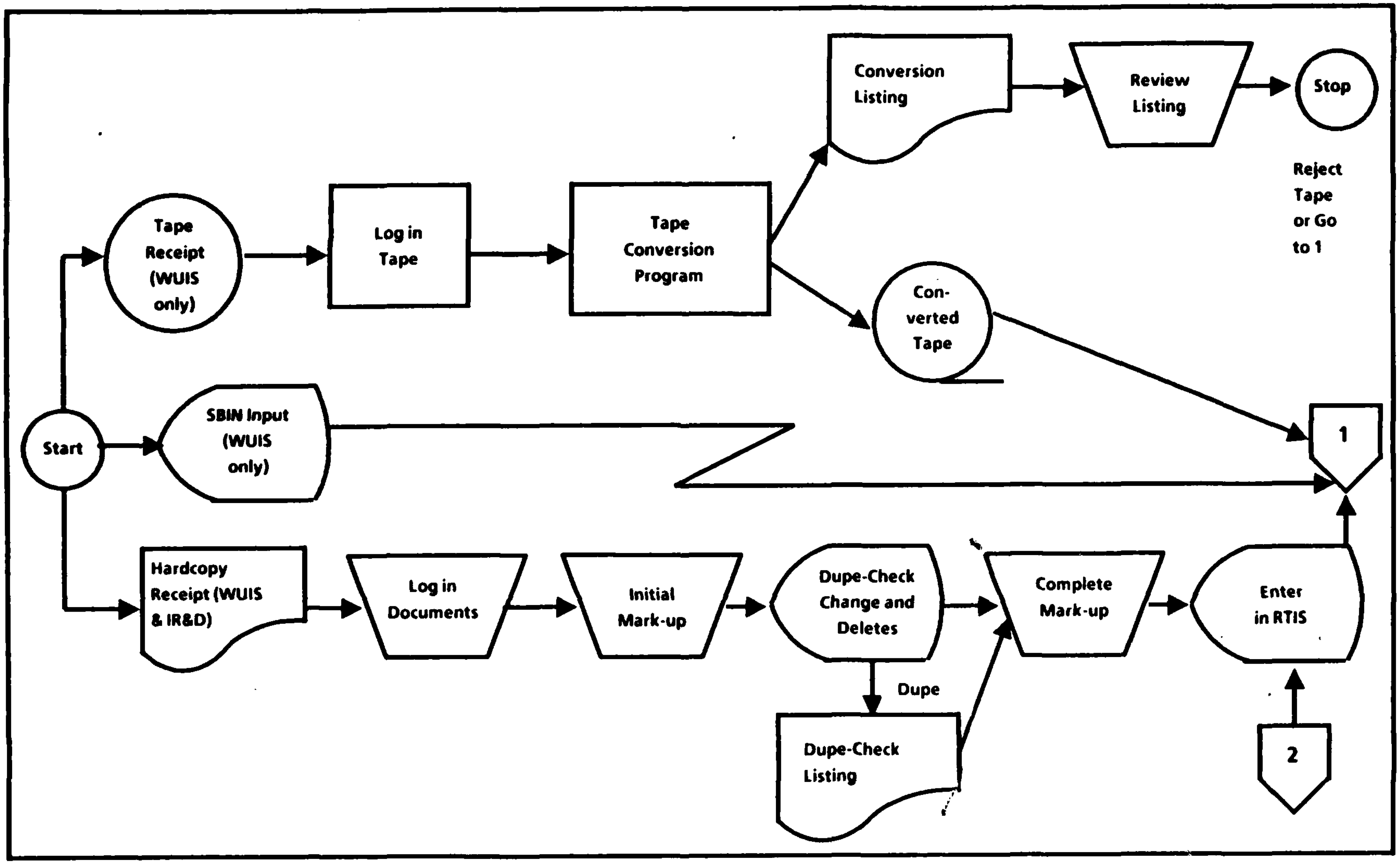
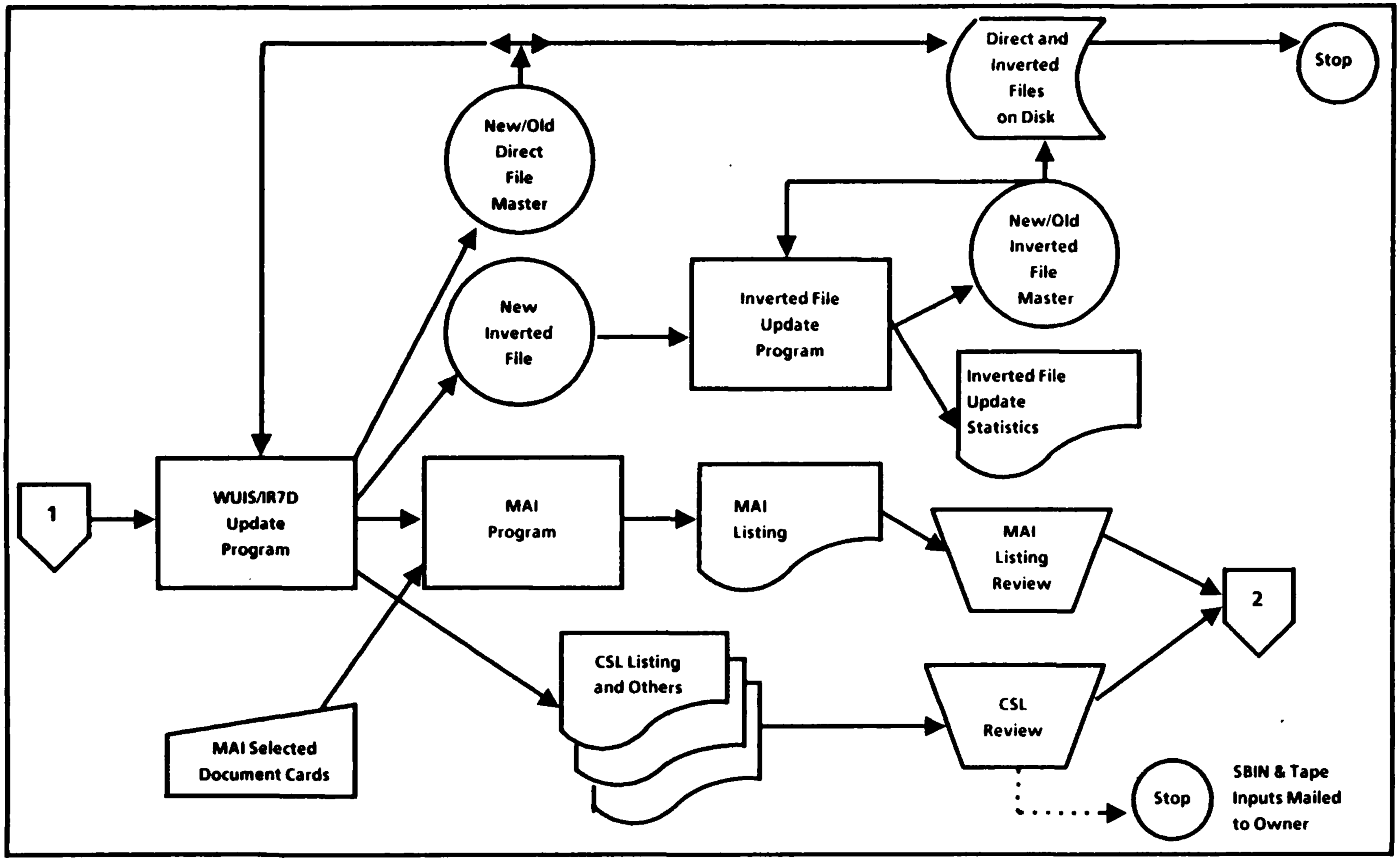


FIGURE D-6. WUIS/IR&D PROCESSING FLOWCHART (CONTINUED)



END
1-87
DTIC